

# SEARCH REQUEST FORM

11-206

Requestor's  
Name: \_\_\_\_\_

Serial  
Number: \_\_\_\_\_

Date: \_\_\_\_\_

Phone: \_\_\_\_\_

Art Unit: \_\_\_\_\_

10607

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 11-17-98  
 Searcher: PMB x8-4291  
 Terminal time: 12  
 Elapsed time: prep 18  
 CPU time: \_\_\_\_\_  
 Total time: \_\_\_\_\_  
 Number of Searches: \_\_\_\_\_  
 Number of Databases: 11

### Search Site

\_\_\_\_ STIC  
 \_\_\_\_ CM-1  
 \_\_\_\_ Pre-S

### Type of Search

0 N.A. Sequence  
2 A.A. Sequence  
 \_\_\_\_ Structure  
 \_\_\_\_ Bibliographic

### Vendors

\_\_\_\_ IG  
 \_\_\_\_ STN  
 \_\_\_\_ Dialog  
 \_\_\_\_ APS  
 \_\_\_\_ Geninfo  
 \_\_\_\_ SDC  
 \_\_\_\_ DARC/Questel  
113 Other



\*\*\*\*\*

W P S R E H

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri Nov 13 22:12:04 1998; MasPar time 1501.65 Seconds  
Tabular output not generated. 1522.943 Million cell updates/sec

Title: >US-08-887-977-9

Description: (1-1119) from US0887977.seq

Perfect Score: 1119

N.A. Sequence: 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCCCTCAA 1119

Comp: TACAAAGCTGAGTGCACTT.....TCTTTCGACTCAGAGGGATT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 552174 seqs, 1021863385 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb155

1:em\_ba 2:em\_htg 3:em\_hum1 4:em\_hum2 5:em\_in 6:em\_om  
7:em\_or 8:em\_ov 9:em\_pat 10:em\_pl 11:em\_ro

Database: genbank107  
12:gb\_ba 13:gb\_htg 14:gb\_in 15:gb\_om 16:gb\_ov 17:gb\_pat  
18:gb\_ph 19:gb\_pl 20:gb\_pr1 21:gb\_pr2 22:gb\_ro 23:gb\_st  
24:gb\_sts 25:gb\_sy 26:gb\_un 27:gb\_v1

Statistics: Mean 10.987; Variance 4.759; scale 2.308

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1104	98.7	1255	21	HSCKRL3	H. sapiens G protein-co	0.00e+00
2	1104	98.7	3693	21	HSU45984	Homo sapiens CCR6 chem	0.00e+00
3	1100	98.3	1518	21	HSU68032	Human G protein-couple	0.00e+00
4	1100	98.3	2978	21	HSU68030	Human G protein-couple	0.00e+00
5	1070	95.6	1137	20	HSU60000	Human IL8-related rece	0.00e+00
6	530	47.4	1302	22	AB009369	Mus musculus mRNA for	0.00e+00
7	75	6.7	1500	20	HSDNABLR2	H. sapiens BLR2 gene.	1.42e-39
8	75	6.7	2139	20	HUMEBII1CDN	Human G protein-couple	1.42e-39
9	75	6.7	2154	20	HUMGPCRA	Human Epstein-Barr vir	1.42e-39
10	75	6.7	2215	20	HUMEBII103	Human G protein-couple	1.42e-39
11	75	6.7	133801	21	AC004585	Homo sapiens chromoso	1.42e-39
12	64	5.7	2577	20	HSU45982	Human G protein-couple	3.27e-30
13	61	5.5	1495	20	HUMCCCRA1A	Human C-C chemokine re	1.04e-27
14	61	5.5	1609	17	E13385	CDNA encoding human MI	1.04e-27
15	62	5.5	2072	22	MUSEBII1CDN	Mouse G protein-couple	1.54e-28

16	61	5.5	2156	17	I58541	Sequence 1 from patent	1.04e-27
17	61	5.5	2156	20	HUMRANTES	Homo sapiens macrophag	1.04e-27
18	61	5.5	2214	21	HUMHML145	Human mRNA for HM145.	1.04e-27
19	59	5.3	528	22	MMLCR11	Mus musculus mRNA enco	4.72e-26
20	59	5.3	1180	22	MMLCR12	Mus musculus lcr-1 gen	4.72e-26
21	59	5.3	1223	22	MMU59760	Mus musculus fusin hom	4.72e-26
22	59	5.3	1422	22	MMLCR13	Mus musculus lcr-1 gen	4.72e-26
23	59	5.3	1877	22	D87747	Mouse mRNA for murine	4.72e-26
24	59	5.3	3366	22	MMU65580	Mus musculus fusin (CX	4.72e-26
25	58	5.2	1010	22	RNU90610	Rattus norvegicus CX	3.14e-25
26	58	5.2	7218	17	I66494	Sequence 14 from paten	3.14e-25
27	57	5.1	1068	15	RAB1L8C	Oryctolagus cuniculus	2.08e-24
28	57	5.1	1068	21	AF017282	Macaca mulatta chemoki	2.08e-24
29	57	5.1	1232	15	RAB1L8REC	Oryctolagus cuniculus	2.08e-24
30	57	5.1	1575	22	MMLSTRPT	M. musculus mRNA for le	2.08e-24
31	57	5.1	1656	16	AB010713	Cyprinus carpio mRNA f	2.08e-24
32	57	5.1	1744	15	BTU19947	Bos taurus interleukin	2.08e-24
33	57	5.1	1809	22	AB000803	Mouse mRNA for murine	2.08e-24
34	57	5.1	3770	22	MMLSTRGN	M. musculus gene encodi	2.08e-24
35	55	4.9	1068	21	AF026535	Homo sapiens chemokine	8.88e-23
36	55	4.9	1201	20	HSU28694	Human eosinophil CC ch	8.88e-23
37	55	4.9	1689	20	HSU49727	Human C-C chemokine re	8.88e-23
38	55	4.9	1717	21	HSU51241	Human eosinophil eotax	8.88e-23
39	53	4.7	1068	21	MMV13776	Macaca mulatta CCR-3 g	3.68e-21
40	53	4.7	1068	21	AF017283	Macaca mulatta chemoki	3.68e-21
41	51	4.6	1083	17	E13909	CDNA encoding human MC	1.47e-19
42	51	4.6	1083	21	HUMMCP1R	Human mRNA for monocy	1.47e-19
43	52	4.6	1140	15	FCU92795	Felis catus fusin (CXC	2.34e-20
44	51	4.6	1979	20	HSU03905	Human monocyte chemoat	1.47e-19
45	51	4.6	2232	20	HSU03882	Human monocyte chemoat	1.47e-19

ALIGNMENTS

RESULT 1

LOCUS	HSCKRL3	1255 bp	DNA	PRI	13-NOV-1996
DEFINITION	H.sapiens G protein-coupled receptor	CCR-L3.			
ACCESSION	Z79784				
NID	gl668737				
KEYWORDS	G Protein-coupled Receptor CCR-L3.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	Gutierrez, J., Varona, R., Zaballos, A., Lind, P. and Marquez, G.				
JOURNAL	Unpublished				
REFERENCE	2	(bases 1 to 1255)			
AUTHORS	Zaballos, A.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-SEP-1996) Angel Zaballos, Research, Pharmacia & Upjohn, Antonio Lopez 109, Madrid, 28026, Spain				
REFERENCE	3	(bases 1 to 1255)			
AUTHORS	Zaballos, A., Varona, R., Gutierrez, J., Lind, P. and Marquez, G.				
TITLE	Molecular cloning and RNA expression of two new human chemokine receptor-like genes				
JOURNAL	Biochem. Biophys. Res. Commun. 227 (3), 846-853 (1996)				
MEDLINE	97040707				
FEATURES	Location/Qualifiers				
source	1..1255				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="6"				
	76..1185				
	/codon_start=1				
	/product="G PROTEIN-COUPLED RECEPTOR CCR-L3"				
	/db_xref="PID:e264774"				
	/db_xref="PID:g1668738"				
	/translation="MNFSDVDSSEDFVSVNTSYVSVDSEMLLCSLQEVRFSLRVF				
	PLAYSLLICVFLGLNIIIVITFAFKKRSMTDYLLNMAIDLEVLFWLPAWVASHA				
	TGAWFSNATCKLKGIIYAINFCNGLMLTCTISMDRYIAIVQATKRSRLRSLPRSK				
	IICLVVWGLSVIISSTFTVENQKNTQSGDCEPKYQTVSEPIRWKLLMLGLELFLFGF				



FIPLMFICYFTYFIVKTLVQAKSRHKAIRVIAVVLVFLACQIPHNMYLVTAANL  
GKNRSQSEKLLGYTFTVTEVLAFLHCCNLNPLYAFIGKFRNYFLKILKDLWCVR  
KYASSGFCAGRISENISROTSETADNDASSFTM

BASE COUNT 302 a 284 g 375 t  
ORIGIN

Query Match 98.7%; Score 1104; DB 21; Length 1255;  
Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
Matches 1114; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Db 89 ATGTTTCGACTCCAGTGAAGATATTTTGTCTGAGTCAATACATCTCATATTACTCAGTTG 148  
QY 1 ATGTTTCGACTCCAGTGAAGATATTTTGTCTGAGTCAATACATCTCATATTACTCAGTTG 60  
Db 149 ATTCAGATGTTTACTGTCCTCTTGAGGAGGTCAGGAGTCTCTCCAGGCTATTTGTAC 208  
QY 61 ATTCAGATGTTTACTGTCCTCTTGAGGAGGTCAGGAGTCTCTCCAGGCTATTTGTAC 120  
Db 209 CGA-TTGCCCTACTCCTTGATCTGTCTTTGCGCTCCTGGGGAATATTTCTGGTGTATC 267  
QY 121 CGAATGCCCTACTCCTTGATCTGTCTTTGCGCTCCTGGGGAATATTTCTGGTGTATC 180  
Db 268 ACCTTTGCTTTTATAGAAAGCCAGGCTCTATGACAGACGCTCTATCTTGAACATGGCC 327  
QY 181 ACCTTTGCTTTTATAGAAAGCCAGGCTCTATGACAGACGCTCTATCTTGAACATGGCC 240  
Db 328 ATTGCAGATCCTCTTGTGTCTTACTCTCCATCTCTGGGAGTGAGTCATGCCACTGGT 387  
QY 241 ATTGCAGATCCTCTTGTGTCTTACTCTCCATCTCTGGGAGTGAGTCATGCCACTGGT 300  
Db 388 GCCTGGGTTTTCAGCAATGCCAGTCAAGTGTCTTAAAGCATCTATGCCATCACTTT 447  
QY 301 GCCTGGGTTTTCAGCAATGCCAGTCAAGTGTCTTAAAGCATCTATGCCATCACTTT 360  
Db 448 AACTGGGGATGCTGCTCCTGACTGTGATGACGACGCTATGACGACGCTATGACGACGCTATGACG 507  
QY 361 AACTGGGGATGCTGCTCCTGACTGTGATGACGACGCTATGACGACGCTATGACGACGCTATGACG 420  
Db 508 GCGACTAAGTCATTCGCGCTCCGATCCAGACACTACCGCGCAGCAAAATCATCTGCCTT 567  
QY 421 GCGACTAAGTCATTCGCGCTCCGATCCAGACACTACCGCGCAGCAAAATCATCTGCCTT 480  
Db 568 GTTGTGGGGGCTGCAGTCATCATCTCAGCTCAACTTTTGTCTTCAACCAAAATAC 627  
QY 481 GTTGTGGGGGCTGCAGTCATCATCTCAGCTCAACTTTTGTCTTCAACCAAAATAC 540  
Db 628 AACACCAAGGACGATGCTGTGAACCAAGTACGAGCTGTCTCGGAGCCCATCAGG 687  
QY 541 AACACCAAGGACGATGCTGTGAACCAAGTACGAGCTGTCTCGGAGCCCATCAGG 600  
Db 688 TGAAGCTGCTGATGTTGGGCTTGAGCTACTCTTTGTTTCTTTATCCCTTTGATGTTT 747  
QY 601 TGAAGCTGCTGATGTTGGGCTTGAGCTACTCTTTGTTTCTTTATCCCTTTGATGTTT 660  
Db 748 ATGATATTTTGTACAGTTTCATGTCATAAACCCTTGTGCAAGTCAAGATCTTAAAGG 807  
QY 661 ATGATATTTTGTACAGTTTCATGTCATAAACCCTTGTGCAAGTCAAGATCTTAAAGG 720  
Db 808 CACAAGCCATCCGTGATCATGCTGTGGTGTGTTGTTCTTCTTCTGCTTGTGCAATTCCT 867  
QY 721 CACAAGCCATCCGTGATCATGCTGTGGTGTGTTGTTCTTCTTCTGCTTGTGCAATTCCT 780  
Db 868 CATAACATGCTGCTGTTGAGGCTGCAAAATTTGGTGAATGAACCGATCTGCGCAG 927  
QY 781 CATAACATGCTGCTGTTGAGGCTGCAAAATTTGGTGAATGAACCGATCTGCGCAG 840  
Db 928 AGCGAAAGCTAATTTGGCTATACGAAATCTGTCAGAAAGTCTGCTGCTTCTGCTGCTG 987  
QY 841 AGCGAAAGCTAATTTGGCTATACGAAATCTGTCAGAAAGTCTGCTGCTTCTGCTGCTG 900  
Db 988 TGCCTGAACCCCTGCTCTACGCTTTTATTTGGGAGAGTTCAGAAACTACTTTCTGAAG 1047  
QY 901 TGCCTGAACCCCTGCTCTACGCTTTTATTTGGGAGAGTTCAGAAACTACTTTCTGAAG 960

Db 1048 ATCTTGAAGGACCTGTGTGTGTGTGAGAGGAAGTACAAAGTCTCTCAGGCTTCTCCTGTGCC 1107  
QY 961 ATCTTGAAGGACCTGTGTGTGTGTGAGAGGAAGTACAAAGTCTCTCAGGCTTCTCCTGTGCC 1020  
Db 1108 GGGAGGTACTCAGAAAACATTTCTGGCAGACCAAGTGTGAGCCGACAGATACCAATGCG 1167  
QY 1021 GGGAGGTACTCAGAAAACATTTCTGGCAGACCAAGTGTGAGCCGACAGATACCAATGCG 1080  
Db 1168 TCGTCTTCTCAGTATGTATAGAAAGCTGAGTCTCCCTAA 1206  
QY 1081 TCGTCTTCTCAGTATGTATAGAAAGCTGAGTCTCCCTAA 1119

RESULT 2  
LOCUS HSU45984 3693 bp DNA PRI 09-JUL-1997  
DEFINITION Homo sapiens CCR6 chemokine receptor (CMKR6) gene, complete cds.  
ACCESSION U45984  
NID 92246432  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
Homo.

REFERENCE 1 (bases 1 to 3693)  
AUTHORS Baba, M., Imai, T., Nishimura, M., Kakizaki, M., Takagi, S.,  
Hieshima, K., Nomiya, H. and Yoshie, O.  
TITLE Identification of CCR6, the specific receptor for a novel  
lymphocyte-directed CC chemokine LARC  
J. Biol. Chem. 272 (23), 14893-14898 (1997)

REFERENCE 2 (bases 1 to 3693)  
AUTHORS Lautens, L.L., Modi, W. and Bonner, T.I.  
TITLE Cloning, Tissue Distribution and Chromosomal Localization of a  
Potential G-Protein-Linked Receptor  
Unpublished

REFERENCE 3 (bases 1 to 3693)  
AUTHORS Bonner, T.I.  
TITLE Direct Submission  
JOURNAL Submitted (18-JAN-1996) Tom I. Bonner, Lab of Cell Biology, NIMH,  
Bldg 36, Rm 3A-17, MSC 4090, Bethesda, MD 20892-4090, USA

FEATURES  
source  
1. 3693  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="6q27"  
/clone="GPR-CY4"  
join(35..140,237..3136)  
132..1352  
/gene="CMKR6"  
join(132..140,237..1352)  
/gene="CMKR6"  
/note="G protein-coupled receptor"  
/product="CCR6 chemokine receptor"  
/db\_xref="PID:92251211"

mRNA  
gene  
CDS

polyA\_signal  
polyA\_site  
BASE COUNT 1015 a 726 c 828 g 1124 t

ORIGIN

Query Match 98.7%; Score 1104; DB 21; Length 3693;  
Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
Matches 1114; Conservative 0; Mismatches 4; Indels 1; Gaps 1;



Db 1335 TCGTCCCTCACTATGTATAGAGAGCTGAGTCCCTAA 1373  
|||||  
QY 1081 TCGTCCCTCACTATGTATAGAGAGCTGAGTCCCTAA 1119  
|||||

RESULT 3  
LOCUS HSU68032 1518 bp DNA PRI 06-MAR-1997  
DEFINITION Human G protein-coupled receptor (STRL22) gene, complete cds.  
ACCESSION U68032  
NID g1870668  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
AUTHORS Liao, F., Lee, H. H. and Farber, J. M.  
TITLE Cloning of STRL22, a new human gene encoding a G-protein-coupled  
receptor related to chemokine receptors and located on chromosome  
6q27  
JOURNAL Genomics 40 (1), 175-180 (1997)  
MEDLINE 97224503  
REFERENCE 2 (bases 1 to 1518)  
AUTHORS Farber, J. M. and Liao, F.  
TITLE Direct Submission  
JOURNAL Submitted (26-AUG-1996) NIAID/Bldg10/Rm11N-228, NIH, 9000 Rockville  
Pike, Bethesda, MD 20892, USA  
FEATURES  
source  
1. .1518  
/organism="Homo sapiens"  
/note="Caucasian"  
/db\_xref="taxon:9606"  
/clone\_lib="Stratagene # 951202"  
/chromosome="6"  
/map="6q27"  
/sex="male"  
1. .1518  
/gene="STRL22"  
intron <1. .2  
/gene="STRL22"  
exon 3. .108  
/gene="STRL22"  
join(100. .108,205. .1320)  
CDS /gene="STRL22"  
/note="contains seven transmembrane domains; expressed in  
lymphocytes"  
/codon\_start=1  
/product="G protein-coupled receptor"  
/db\_xref="PID:g1870669"  
/translation="MSGSMNFDVDFDSSDYFVSNTSYSDSEMLCSLQEVRF  
SLFVPIAYSLICVGLGNLIVITFAFYKARSMTDVYLLNMAIDILFVLTLPFW  
AYSHATGAWFVSNTCKLLGIYAINFCGMILLTICISMDRYIAIQAATKSFRLSRT  
LPRTKILIVWGLSVIISSSFVFNQKYNQSDVCEPKQIVSEPRKWLMLGLE  
LIFGFFIPLEMFICYTFIVKTLVQNSKRRAIRVIAVLVFLACQIPNMLLV  
TAANGLMNRSCSEKLYGTIVTEVLAFLHCLNPLVYAFYQKFRNYFLKILKDL  
WCVRRYKSGSFGSCAGRYSENISROTSETADNDNASSFTN"  
109. .204  
intron  
exon 205. .>1518  
/gene="STRL22"  
BASE COUNT 360 a 354 c 351 g 453 t  
ORIGIN  
Query Match 98.3%; Score 1100; DB 21; Length 1518;  
Best Local Similarity 99.4%; Pred. No. 0.00e+00;  
Matches 1112; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
Db 224 ATGTTTCGACTCCAGTGAAGATTATTTGTGTCAGTCAATACCTCATATCTACCTG 283  
|||||  
QY 1 ATGTTTCGACTCCAGTGAAGATTATTTGTGTCAGTCAATACCTCATATCTACCTG 60  
|||||

Db 284 ATCTCAGATGTTACTGTCTGCTTCCTTCGAGGAGGTCAGGAGTCTCCAGGCTATTGTAC 343  
|||||

Db 256 ATGTTTCGACTCCAGTGAAGATTATTTGTGTCAGTCAATACCTCATATCTACCTG 315  
|||||  
QY 1 ATGTTTCGACTCCAGTGAAGATTATTTGTGTCAGTCAATACCTCATATCTACCTG 60  
|||||

Db 316 ATCTGAGATCTTACTGTCTCCTTCGAGGAGTCCAGGAGTCTCCAGGCTATTTGTAC 375  
|||||  
QY 61 ATCTGAGATCTTACTGTCTCCTTCGAGGAGTCCAGGAGTCTCCAGGCTATTTGTAC 120  
|||||

Db 376 CGA-TTGGCTACTCTTGATCTGTCTTTGGCCCTCTGGGGAATATCTGTGGTGAATC 434  
|||||  
QY 121 CGAATTCCTACTCTTGATCTGTCTTTGGCCCTCTGGGGAATATCTGTGGTGAATC 180  
|||||

Db 435 ACCTTTGTCTTTTAAAGAGCCAGGCTATGACAGAGCTCTATCTTTGACATGGCC 494  
|||||  
QY 181 ACCTTTGTCTTTTAAAGAGCCAGGCTATGACAGAGCTCTATCTTTGACATGGCC 240  
|||||

Db 495 ATTCGAGACATCTCTTTTGTCTTACTCTCCCATCTTGGGCAGTGAAGTCAATGCCACTGGT 554  
|||||  
QY 241 ATTCGAGACATCTCTTTTGTCTTACTCTCCCATCTTGGGCAGTGAAGTCAATGCCACTGGT 300  
|||||

Db 555 GCGTGGGTTTTCAGAAATGCCAGCTGCAAGTTGCTTAAAGGCAATCTATGCCATCAACTTT 614  
|||||  
QY 301 GCGTGGGTTTTCAGAAATGCCAGCTGCAAGTTGCTTAAAGGCAATCTATGCCATCAACTTT 360  
|||||

Db 615 AACTGCGGGATGCTCTCTGCTGATCATCTCCAGCTCAACTTTTGTCTCAACCAAAATAC 674  
|||||  
QY 361 AACTGCGGGATGCTCTCTGCTGATCATCTCCAGCTCAACTTTTGTCTCAACCAAAATAC 420  
|||||

Db 675 GCGACTAAGTCAATCCGGCTCCGATCCAGAACACTACCCGAGCAAAATCATCTCGCTT 734  
|||||  
QY 421 GCGACTAAGTCAATCCGGCTCCGATCCAGAACACTACCCGAGCAAAATCATCTCGCTT 480  
|||||

Db 735 GTTGTGTGGGGCTGTCACTCATCTCCAGCTCAACTTTTGTCTCAACCAAAATAC 794  
|||||  
QY 481 GTTGTGTGGGGCTGTCACTCATCTCCAGCTCAACTTTTGTCTCAACCAAAATAC 540  
|||||

Db 795 AACACCCAGGAGGATGTCTGTGAACCAAGTACCAGACTGTCTGGAGCCCATCAG 854  
|||||  
QY 541 AACACCCAGGAGGATGTCTGTGAACCAAGTACCAGACTGTCTGGAGCCCATCAG 600  
|||||

Db 855 TGGAGCTGTCTGTGGGGCTTCAGTCTACTTTGGTTCTTTATCCCTTTGATGTTTC 914  
|||||  
QY 601 TGGAGCTGTCTGTGGGGCTTCAGTCTACTTTGGTTCTTTATCCCTTTGATGTTTC 660  
|||||

Db 915 ATGATATTTTGTACACGTTTCATTTGCAAAACCTTGGTCAAGCTCAGAAATCTTAAAGG 974  
|||||  
QY 661 ATGATATTTTGTACACGTTTCATTTGCAAAACCTTGGTCAAGCTCAGAAATCTTAAAGG 720  
|||||

Db 975 CACAAAGCCATCCGTTGTAATCATAGCTGTGGTGTCTTTCTGGCTGTCTGATTCCT 1034  
|||||  
QY 721 CACAAAGCCATCCGTTGTAATCATAGCTGTGGTGTCTTTCTGGCTGTCTGATTCCT 780  
|||||

Db 1035 CATACATGGTCCCTGTGTGACGGCTCAAAATTTGGTAAATGAACGATCTCCGAC 1094  
|||||  
QY 781 CATACATGGTCCCTGTGTGACGGCTCAAAATTTGGTAAATGAACGATCTCCGAC 840  
|||||

Db 1095 AGCGAAAGCTAAATGGCTATACGAAACTGTCAAGAACTCTGGCTTTCTGCACTGC 1154  
|||||  
QY 841 AGCGAAAGCTAAATGGCTATACGAAACTGTCAAGAACTCTGGCTTTCTGCACTGC 900  
|||||

Db 1155 TGCCTGAACCTGTCTACGCTTTTATTTGGGCGAGAAGTTTCAGAACTACATTTCTGAAG 1214  
|||||  
QY 901 TGCCTGAACCTGTCTACGCTTTTATTTGGGCGAGAAGTTTCAGAACTACATTTCTGAAG 960  
|||||

Db 1215 ATCTTGAAGGACCTGTGTGTGTGAGAGGAAGTCAAGTCTCTAGGCTTCTCTGTGCC 1274  
|||||  
QY 961 ATCTTGAAGGACCTGTGTGTGTGAGAGGAAGTCAAGTCTCTAGGCTTCTCTGTGCC 1020  
|||||

Db 1275 GGGAGGTACTCAGAAACATTTCTCGGAGACCACTGAGACCGCAGATACAGCAATGGG 1334  
|||||  
QY 1021 GGGAGGTACTCAGAAACATTTCTCGGAGACCACTGAGACCGCAGATACAGCAATGGG 1080  
|||||











```

688 ATGATATTTTGTACGTTTCATTTGTCACAAACCTTGGTGCAGCTCAGAAATCTTAAAGG 747
|||||
561 ATGATATTTTGTACGTTTCATTTGTCACAAACCTTGGTGCAGCTCAGAAATCTTAAAGG 720
|||||
748 CACAAAGCCATCCGTGTAATCATAGCTGTGGTGTGTTTCTGGCTTGTGAGATTCCT 807
|||||
721 CACAAAGCCATCCGTGTAATCATAGCTGTGGTGTGTTTCTGGCTTGTGAGATTCCT 780
|||||
808 CATAACATGTCCTGCTGTGACGGCTGCAAAATTTGGTGAATAAGAACCGATCTGCCAG 867
|||||
781 CATAACATGTCCTGCTGTGACGGCTGCTAATTTGGTGAATAAGAACCGATCTGCCAG 840
|||||
868 AGCGAAAGCTAATTTGCTATACAAACCTGTCAGAGTCTTGGCTTCTGCTGACATGC 927
|||||
841 AGCGAAAGCTAATTTGCTATACAAACCTGTCAGAGTCTTGGCTTCTGCTGACATGC 900
|||||
928 TGCTGAACCTGTGCTCTAGCTTTTATTTGGGAGAAAGTTCAGAAATCTTCTGAAG 987
|||||
901 TGCTGAACCTGTGCTCTAGCTTTTATTTGGGAGAAAGTTCAGAAATCTTCTGAAG 960
|||||
988 ATCTTGAAGACCTGTGGTGTGAGAGGAAGTACAAAGTCTTCTGCTGCTGCC 1047
|||||
961 ATCTTGAAGACCTGTGGTGTGAGAGGAAGTACAAAGTCTTCTGCTGCTGCC 1020
|||||
1048 GCGAGTACTCAGAAACATTTCTCGCAGACCACTGAGAGCGGAGATACAGCAATGCC 1107
|||||
1021 GCGAGTACTCAGAAACATTTCTCGCAGACCACTGAGAGCGGAGATACAGCAATGCC- 1079
|||||
1108 GTGCTCTTCACTATGTATAGAAAGCTGA 1137
|||||
1080 GTGCTCTTCACTATGTATAGAAAGCTGA 1109
|||||

RESULT 6
LOCUS AB009369 1302 bp mRNA ROD 09-DEC-1997
DEFINITION Mus musculus mRNA for G protein-coupled receptor KY411, complete cds.
ACCESSION AB009369
NID 92668409
KEYWORDS G protein-coupled receptor KY411.
SOURCE Mus musculus CDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1302)
Yanagihara,S., Komura,E. and Yamaguchi,Y.
Mouse G protein-coupled receptor KY411.
Published Only in DataBase (1997) in press
2 (bases 1 to 1302)
Yanagihara,S., Komura,E. and Yamaguchi,Y.
Direct Submission
Submitted (01-DEC-1997) to the DDBJ/EMBL/GenBank databases.
Shigehiro Yanagihara, Kirin Brewery Co., Ltd., Pharmaceutical
Research Laboratory, 3 Miyahara-cho, Takasaki, Gunma 370-12, Japan
(E-mail:syanagihara@kirin.co.jp, tel:81-273-46-9826)
Location/Qualifiers
1. .1302
/organism="Mus musculus"
/db_xref="taxon:10090"
147. .1250
/codon_start=1
/product="G protein-coupled receptor KY411"
/db_xref="PID:d1024681"
/db_xref="PID:g2668410"
/translation="MNSTSEYFGTDDYDNTVEYSIPDHGPCSLIEEVRNTEKVPVPIA
YSLICVGLLGNMVMYAFYKKARSDWDVYLINMAIDILFVLTPFWATHTATNT
WYFSDALCKMLGYAVNFCNGLLACISMDRYIAIVQATKSFVRVSRLLTHSKVIC
VAWVFSIIISPTFFNKKYEQDRCVPRYSRVSPTTWKLLGMLGLELFFGFPP
LLFMVFCYLIKTLVQAQNSKKRRAIRVIAVLVFLACQIPHNVLVLTAVNTGV
GRSCTEKVLATRNVAEVLAFHCLCNPLVLYAFIGQKPRNYFMKIMKDWCMRRKNK
MPGFLCARYSESYISRQSETVENASFTM"
BASE COUNT 314 a 331 c 320 g 337 t
ORIGIN

```

```

Query Match 47.4%; Score 530; DB 22; Length 1302;
Best Local Similarity 76.9%; Pred. No. 0.00e+00;
Matches 786; Conservative 0; Mismatches 232; Indels 4; Gaps 3;

Db 228 TGCTCCCTAGAGAGGTGACAGAACTTCACCAAGGTATTTGTGCG-AATTGCGCTACTCCTT 286
|||||
QY 78 TGCTCCTTCGAGAGGTGAGGAGTTCTCCAGGCTATTTGTACCGAATGCGCTACTCCTT 137
|||||
Db 287 AATATGTGCTTTTGGCTCCTTGGGCAACATTTATGGTGGTGATGACCTTTGGCTTTTACAA 346
|||||
QY 138 GATCTGTGTTTGGCTCCTTGGGGAATATCTGGTGGTGATCACCTTTGCTTTTATAA 197
|||||
Db 347 GAAAGCCAGATCCATGACGTCTACCTGTGTAACATGGCCATCACAGACATCTT 406
|||||
QY 198 GAAGCGAGGTCTATACAGAGGTCTATCTTTGAACATGGCCATTCACACATCTCTT 257
|||||
Db 407 TGCTCCTCAACCTACCGTCTTGGGCAAGTACTCATGCCCAACACATTGGTGGTTCAGCGA 466
|||||
QY 258 TGTCTTCTACTCTCCCATCTTGGGCAAGTACTCATGCCCAACACATTGGTGGTTCAGCGA 317
|||||
Db 467 TGCACTGTGTAACACTGATGAAAGGCAACATATGCGGTCAACTTTAACTGTGGGATGCTGCT 526
|||||
QY 318 TGCCACGTGCAAGTTGCTAAAGGCACTCTATGCCATCACTTTAACTGCGGATGCTGCT 377
|||||
Db 527 CTTGGCTGTATCAGATGGACCGGTACATGTCATGTCAGAGCCATCAGTGAAGTGTGGGTAT 586
|||||
QY 378 CTTGACTTGCATAGCATGGACCGGTACATGTCATGTCAGAGCCATCAGTGAAGTGTGGGTAT 437
|||||
Db 587 GGTAGGCTCCAGAACACATGACGACAGTAAGTCTCTGTGGCAGTGTGGTTCATCTC 646
|||||
QY 438 GCTCCGATCCAGAACACATGACGCGGAGCAAAATCATCTGCCCTTGTGTGGGGGTGTC 497
|||||
Db 647 CATCATCTCTCAAGCCCTACATTTATCTTCAACAGAAATACGAGCTGCAGGATCGTGA 706
|||||
QY 498 AGTCATCATCTCCAGCTCAACTTTGTCTTCAACCAAAATACACACCCAGGACGCA 557
|||||
Db 707 TGTCTGTAGCCACCGGTACAGGTCTGTCTCAGAGCCCATCAGTGAAGTGTGGGTAT 766
|||||
QY 558 TGTCTGTGAACCAAGTACCAAACTGTCTCGGAGCCCATCAGTGAAGTGTGGGTAT 617
|||||
Db 767 GGGACTGGAGCTGTCTTTGGGTCTTACCCTTTGCTTGTGTATGCTTGTCTGCTATCT 826
|||||
QY 618 GGGGCTTGAGCTACTCTTTGGTCTTTTATCCCTTTGATGTTGATATTTTGTATAC 677
|||||
Db 827 GTTCATTTATAGAGCTTGTGTCAGGCCCGAGAACTCCAAGAGCCAGAGCCATCCGAGT 886
|||||
QY 678 GTTCATTTATAGAGCTTGTGTCAGAGCTCAGAAATTTAAAGGCCAAGCCATCCGAGT 737
|||||
Db 887 CTTGATCGCTGTGGTTCGCTGCTTCTGCTGTGTGTCAGATCCCTTCAACATGGTCTCTCT 946
|||||
QY 738 AATCATAGCTGTGGTGTGCTTGTCTTGTGCTTGTGATTTCTCTCAATGCTGCTGCT 797
|||||
Db 947 CTTGACTGGGTTCACACGGGCAAGTGGGCGGAGCTGCGAGCCAGCAAGAAAGTCTCTCG 1006
|||||
QY 798 TGTGACGCTGTCTAATTTGGGTAAATGAACCGATCTCTGCCAGAGCGAAAGTAAATTGG 857
|||||
Db 1007 CTACACGAGGACGTGGGCGGAGTCTGCTGCTTCTGCTGATGCTGCTCAACCCCGTGT 1066
|||||
QY 858 CTATACGAAACTGTACAGAACTCTTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 917
|||||
Db 1067 GTATCGGTTTATGGACAGAAATTCAGAACTACTTTCATGAAGATCATGAAGATGTGTG 1126
|||||
QY 918 CTACGCTTTTATGGGCAAGTTCAGAACTACTTTCAGAGATCTTGAAGGACCTGTG 977
|||||
Db 1127 GTGTATGAGAGAGAAATPAAGATCGCTGCTTCTCTCTGTCGCCGGGTTCATCGGAAAG 1186
|||||
QY 978 GTGTGTGAGAGAAAGTACAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1037
|||||
Db 1187 CTACATCTCCAGGACAGCCAGTGAGACCTCGCAAAATGATATGATGCTGCTCTTACCAT 1246
|||||
QY 1038 C-ATTTC- -GGCAGACCAGTGAGACCGCAGATACAGCAATGCGCTGCTCTTACTAT 1094
|||||

```







```
QY 256 TTTGTTCTTACTCTCCCAATCTGGCGAGTGAGTCATGCCACTGGTGGCTGTTTTCAGC 315
Db 442 GTCCACTTTTGCAGCTCACTTTTGGCCATCTACAAGATGAGCTTCTTCAGTGGCATGCTC 501
QY 316 AATGCCACGTGCAAGTGTCTAAAGGAGCATCTATCCATCACTTTAACTGGCGGATGCTG 375
Db 502 CTACTCTTTTGCATCAGCATGACGCTGATCGTGGCCATCGTCCAGGC 548
QY 376 CTCTGACTTGCATTAGCATGGACCGGTATACATGCCCATTTGTACAGGC 422

RESULT 9 HUMGPCRA 2154 bp mRNA PRI 31-DEC-1994
LOCUS Human Epstein-Barr virus induced G-protein coupled receptor mRNA,
DEFINITION complete cds.
ACCESSION L08176
NID 9183484
KEYWORDS Epstein-Barr virus induced gene; G-protein coupled receptor.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
REFERENCE 1 (bases 1 to 2154)
AUTHORS Birkenbach,M., Josefsen,K., Yalamanchili,R., Lencir,G. and Kieff,E.
TITLE Epstein-Barr virus-induced genes: first lymphocyte-specific G
protein-coupled peptide receptors
J. Virol. 67 (4), 2209-2220 (1993)
93188173
FEATURES
Location/Qualifiers
1..2154
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="BL41/B95-8"
/cell_type="B lymphocyte, EBV-converted Burkitt lymphoma"
/germline
/map="17q12-21.2"
sig_peptide 64..135
/gene="EBI 1"
/product="EBV induced G-protein coupled receptor"
64..1200
/gene="EBI 1"
64..1200
/gene="EBI 1"
/note="putative"
/codon_start=1
/product="EBV induced G-protein coupled receptor"
/db_xref="PID:g183485"
ESLCKKDVRFKAWFLPIMYSICFVGLLNGLVLTIIYFRKLTMTDTYLLNLAV
ADILFLTLFPWAYSAKSWVGFHFKLIFAIYKMSFFSGMLLLCISIDRYVAIQ
AVSAHRARVLLISKLSCVGSAILATVLSIPELLYSDLQSSSEQAMRCSLITEHVE
AFITIQVQMWIGFLVPLILAMSFYLVIRTLQARNFERNKAIKVIAVVVVFVFO
LPYNGVLAOTVANENITSSCELSKQLNIADVTYSLACVRCVNPFLYAFIGVKFR
NDIFKFLDGLCLSQEQLRQSSCRHRRSSMSVEAETITTFSP"
BASE COUNT 486 a 647 c 543 g 478 t
ORIGIN
Query Match 6.7%; Score 75; DB 20; Length 2154;
Best Local Similarity 63.7%; Pred. No. 1.42e-39;
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;
Db 203 TGTGCTCCCAAGAGACGTGCGGAACCTT-TAAAGCCTGGTTCCTCCATCATGACTCTCC 261
QY 76 TGTGCTCCTTGCAGAGGTGACGAGTCTCCAGGCTATTGTACCGAATTCGCTACTCC 135
Db 262 ATCATTTGTTTCGTGGGCTACTGGCAATGGCTGGTCTGTGTGACCTATATCTATTC 321
QY 136 TTGATCTGTGTTCTTGGCCCTCTGGGGAATATTCTGGTGTGATCATCCTTTGCTTTTAT 195
Db 322 AAGAGGCTCAAGACCATGACCGATACCTACTGCTCAACCTGGGGTGGCGAGACATCTC 381
QY 196 AAGAGGCCAGGTCTATGACAGACGTCTATCTCTTGAACATGGCCATTCGACATCTC 255
```

```
Db 382 TTCCTCCTGACCTCCCTTCTGGGCTACAGCGGGCCAA--GTCC-TGGGCTCTCGGT 438
QY 256 TTTGTTCTTACTCTCCCAATCTGGCGAGTGAGTCATGCCACTGGTGGCTGTTTTCAGC 315
Db 439 GTCCACTTTTGCAGCTCACTTTTGGCCATCTACAAGATGAGCTTCTTCAGTGGCATGCTC 498
QY 316 AATGCCACGTGCAAGTGTCTAAAGGAGCATCTATCCATCACTTTAACTGGCGGATGCTG 375
Db 499 CTACTCTTTTGCATCAGCATGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTG 545
QY 376 CTCTGACTTGCATTAGCATGGACCGGTATACATGCCCATTTGTACAGGC 422

RESULT 10 HUMEBI103 2215 bp DNA PRI 10-AUG-1995
LOCUS Human G protein-coupled receptor (EBI 1) gene exon 3, complete cds.
DEFINITION L31584 L31539
ACCESSION
NID 9468314
KEYWORDS G protein-coupled receptor.
SEGMENT 3 of 3
SOURCE Homo sapiens (tissue library: Statagene #946205) placenta DNA.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
REFERENCE 1 (bases 1 to 2215)
AUTHORS Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Schweickart,T.L., Raport,C.J., Godiska,R., Byers,M.G., Eddy,R.L.
Jr., Shows,T.B. and Gray,P.W.
TITLE Cloning of human and mouse EBI1, a lymphoid-specific
G-protein-coupled receptor encoded on human chromosome 17q12-q21.2
Genomics 23 (3), 643-650 (1994)
95154835
FEATURES
Location/Qualifiers
1..2215
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
/tissue_lib="Statagene #946205"
/map="17q12-21.2"
join(L31582:1376..2586,L31583:1..305,1..2215)
/gene="EBI 1"
join(L31582:1439..1448,L31583:221..270,169..1245)
/gene="EBI 1"
/codon_start=1
/product="G protein-coupled receptor"
/db_xref="PID:g468316"
/translocation="MDLGPMSVLVALLVIFQVCLCODEVTDYIGDNTVDYTLF
ESLCKKDVRFKAWFLPIMYSICFVGLLNGLVLTIIYFRKLTMTDTYLLNLAV
ADILFLTLFPWAYSAKSWVGFHFKLIFAIYKMSFFSGMLLLCISIDRYVAIQ
AVSAHRARVLLISKLSCVGSAILATVLSIPELLYSDLQSSSEQAMRCSLITEHVE
AFITIQVQMWIGFLVPLILAMSFYLVIRTLQARNFERNKAIKVIAVVVVFVFO
LPYNGVLAOTVANENITSSCELSKQLNIADVTYSLACVRCVNPFLYAFIGVKFR
NDIFKFLDGLCLSQEQLRQSSCRHRRSSMSVEAETITTFSP"
BASE COUNT 520 a 665 c 545 g 485 t
ORIGIN
Query Match 6.7%; Score 75; DB 20; Length 2215;
Best Local Similarity 63.7%; Pred. No. 1.42e-39;
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;
Db 248 TGTGCTCCCAAGAGACGTGCGGAACCTT-TAAAGCCTGGTTCCTCCATCATGACTCTCC 306
QY 76 TGTGCTCCTTGCAGAGGTGACGAGTCTCCAGGCTATTGTACCGAATTCGCTACTCC 135
Db 307 ATCATTTGTTTCGTGGGCTACTGGCAATGGCTGGTCTGTGTGACCTATATCTATTC 366
QY 136 TTGATCTGTGTTCTTGGCCCTCTGGGGAATATTCTGGTGTGATCATCCTTTGCTTTTAT 195
Db 367 AAGAGGCTCAAGACCATGACCGATACCTACTGCTCAACCTGGCGGTGGCGAGACATCTC 426
```







```
repeat_region complement(18192..18247)
repeat_region /rpt_family="L2"
repeat_region complement(18356..18450)
repeat_region /rpt_family="L2"
repeat_region complement(18539..18649)
repeat_region /rpt_family="MIR"
repeat_region complement(18681..18733)
repeat_region /rpt_family="MIR"
repeat_region 18814..18849
repeat_region /rpt_family="AT-rich"
repeat_region 20228..20478
repeat_region /rpt_family="LMA"
repeat_region complement(22080..22109)
repeat_region /rpt_family="AT-rich"
repeat_region complement(22273..22418)
repeat_region /rpt_family="L2"
repeat_region 22815..22872
repeat_region /rpt_family="L2"
repeat_region 23082..23134
repeat_region /rpt_family="MLR1C"
repeat_region 23136..23419
repeat_region /rpt_family="AluSx"
repeat_region 23430..23562
repeat_region /rpt_family="MLR1C"
repeat_region 23571..23922
repeat_region /rpt_family="L1PA8"
repeat_region 23946..24235
repeat_region /rpt_family="MLR1C"
repeat_region complement(24336..24632)
repeat_region /rpt_family="AluSx"
repeat_region complement(24638..24930)
repeat_region /rpt_family="AluJb"
repeat_region 25050..25139
repeat_region /rpt_family="MER53"
repeat_region complement(25726..25992)
repeat_region /rpt_family="L2"
repeat_region 26383..26903
repeat_region /rpt_family="MLR1E"
repeat_region complement(27486..27718)
repeat_region /rpt_family="L1MB5"
repeat_region 27904..28130
repeat_region /rpt_family="MIR"
repeat_region complement(28396..28442)
repeat_region /rpt_family="(CA)n"
repeat_region complement(28780..29063)
repeat_region /rpt_family="L2"
repeat_region 29391..29484
repeat_region /rpt_family="(CA)n"
repeat_region 29619..29790
repeat_region /rpt_family="MIR"
repeat_region 29825..30118
repeat_region /rpt_family="AluSx"
repeat_region 30269..30459
repeat_region /rpt_family="MIR"
repeat_region complement(30623..30772)
repeat_region /rpt_family="AluSg/x"
repeat_region 32217..32510
repeat_region /rpt_family="AluSg"
repeat_region 32632..32920
repeat_region /rpt_family="AluSg"
repeat_region 32970..33089
repeat_region /rpt_family="MIR"
repeat_region 33777..33984
repeat_region /rpt_family="MIR"
repeat_region complement(33996..34114)
repeat_region /rpt_family="(GAA)n"
repeat_region complement(35345..35437)
repeat_region /rpt_family="MER5A"
repeat_region 36093..36208
repeat_region /rpt_family="(CATG)n"
repeat_region complement(36908..37017)
```

...
Note: remainder of annotations omitted.

```
Query Match 6.7%; Score 75; DB 21; Length 133801;
Best Local Similarity 63.7%; Pred. No. 1.42e-39;
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 4;

Db 50249 GCGTGGAGCATGGCCAGCTAGCGGTCAATGCTGATGCGCAAGAGTAGGAGCATGCCACTG 50308
||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 422 GCGTGTACAATGGCGATGTACCGGTCCATGCTTAATCAAGTCAGGAGCATGCCGCAG 363

Db 50309 AAGAAGCTCATCTTGTAGATGGCAAGAGTAGCTTGCATAAAGTGGACACCGACGCCAG 50368
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Cp 362 TTAAGTTGATGGCATAGATGCTTTTACCAACTTGCACGTCGCATTGCTGAACCCAC 303

Db 50369 G-ATTG-GCCCG-CTGTAGGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 50425
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Cp 302 GCACCAAGTGGCATGACTCACTGCCAGATGGGAGAGTAAGAACAAGAGGAGTGTGCA 243

Db 50426 ACCGCCAGTTGACAGGTAGTATCGGTATCGGTCTTGAGGCTCTTGAATAGATATAG 50485
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Cp 242 ATGGCCATGTTCAAGAGATAGACGCTGTCTATAGACCTGGCCTCTTATAAAGCAAG 183

Db 50486 GTCAACAGCACGCCCATTTGCCAGTAGCCAGCCAGCAACAAATGATGAGTACATGATA 50545
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Cp 182 GTGATCACCAACAGATATTTCCAGAGGAGGAGCAACACATCAGAGTAGGCAAT 123

Db 50546 GGGAGGAACACAGGCTTTA-AGTTCGCCAGCTCTTCTTGGAGCACA 50591
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 122 CGGTACAAATAGCCTGGAGAACTGCTGACCTCCTGCAAGGAGCACA 76
```

```
RESULT 12
LOCUS HSD45982 2577 bp DNA PRI 02-APR-1996
DEFINITION Human G protein-coupled receptor GPR-9-6 gene, complete cds.
ACCESSION U45982
NID 91245054
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2577)
AUTHORS Lautens,L.L., Tiffany,H.L., Gao,J.-L., Modi,W., Murphy,P.M. and Bonner,T.I.
TITLE Cloning, Tissue Distribution and Chromosomal Localization of two potential G-Protein-Linked Chemokine Receptors
JOURNAL Unpublished
AUTHORS Bonner,T.I.
TITLE Direct Submission
JOURNAL Submitted (16-JAN-1996) Tom I. Bonner, Lab of Cell Biology, NIMH, Bldg 36, Rm 3A-17, MSC 4090, Bethesda, MD 20892-4090, USA
FEATURES
Source
1..2577
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p21.3-22"
58..1131
/note="G protein-coupled receptor"
/codon_start=1
/product="GPR-9-6"
/db_xref="PID:91245055"
/translation="MADYGESESTSMEDYVNFNFDFCEKNVROFASHFLPLPLW
LVFIVGALNSLVILVYCTRYKTMDELLNLAIADRLLELVLPFWAIAAOWKF
QTFMCKVNSMKNFYSVLLIMCISVDRIYIAQAMRAHREKRLLYLMKVCFFI
WVLAALCIPILYLSQESGIACTMTPSDESKLSAVLKLVLGFLPFVPM
ACCTTIIILQAKSKKALKAVITVLTVFVLSQFPYCNILIVQIDAYAMFISN
CAVSTNIDICFOVTOTIAFHSCLNPLVYFVGFRFRDLVLTCLKNLGICISQAQWFSF
TRREGSLKLSMLLETTSGSLI"
BASE COUNT 628 a 613 c 574 g 762 t
ORIGIN
```

Query Match 5.7%; Score 64; DB 20; Length 2577;



```
Best Local Similarity 58.2%; Pred. No. 3.27e-30;
Matches 335; Conservative 0; Mismatches 235; Indels 6; Gaps 6;

Db 265 ATGACCGACATCTCTCTTTGATTTGGCAATGCTGACCTCTCTTTCTTGTCACCTCTT 324
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 ATGACAGAGCTATCTCTTTGACATGGCCATTCGACATCATCTCTTTGTTTCTTACTCTC 270
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 CCCTTCTGGCCATT-G-C-TGCTGTCACCAAGTTCAGAGCTTCATGTCGAAG 381
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 CCATCTGGGAGTAGTATGACCATGCTGCTGGGTTTACAGATGCCAGCTGCAAG 330
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 GTGGTCAACAGCATGATGACATGAACTTCTACAGCTGTGTGTTGCTGATCATGTGCATC 441
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 331 TTGCTAAAAGGCATCTATGCCATCAACTTAACTGCGGGATGCTGCTCTGACTTGCAAT 390
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 AGCGTGGACAGTATGATGCCATGCGCCAGGCGCATGAGACCATATTTGGAGGAGAAA 501
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 391 AGCATGGGCGGTATGATCGCCATTTACAGCGCACTAAAGTCAATTCGCGCTCCGATCCAGA 450
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 AGGCTTTTCTACAGCAAAATGTTTCTTTTACCATCTGGTATTGGCAGCTGCTCTCTGC 561
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 451 ACACATACCGCGCAGCAAAATCACTGCTGTTGTTGTTGGGGGCTGTCAGTCAATCATCTCC 510
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 ATCCAGAAATCTTATACAGCAAAATCAAGAGGAATCCGGCATTTGCTATCTGCACCAATG 621
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 511 AGCTCAACTTTGCTCTCAACCAAAATACACACCCAGGAGCGATGCTGTGACCC 570
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 GTTTACCCCTAGCGA-TGAGAGCACCAAAAC-TGAAGTCAGCTG-TCTTGACCCCTGAAGTGC 678
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 571 AAGTACCAAACTGCTCTCGGAGCCATCAGGTGGAAGTGTGATGTTGGGGCTTGAGCTA 630
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 ATTCCTGGGCTTCTCTCTCCCTCGTGGTATGCTGCTGCTATACCATCATTCATTCAC 738
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 631 CFTCTTGGTCTTCTTATCCCTTTGATGTTTCATGATATTTTGTACAGCTTCATTTGTCAAA 690
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 ACCCTGATACAAAGCAAGTCTTCCAGACACAAAGCCCTAAAGTGACCATCAGCTGTC 798
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 691 ACCTTGGTCAAGCTCAGATTTCTAAAGGACAAAGCCATCCGCTGTAATCATAGCTGTG 750
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 CTGACCGCTTTTGTCTGTCTCAGTTTCCCTACAAAC 834
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 751 GTGCTGTGTTTCTGCTGCTGATTCCTCTCATAAC 786
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13 HUMCCCKR1A 1495 bp mRNA PRI 31-DEC-1994
LOCUS Human C-C chemokine receptor type 1 (C-C CKR-1) mRNA, complete cds.
DEFINITION
ACCESSION L09230
NID 9179984
KEYWORDS C-C chemokine receptor type 1.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1495)
AUTHORS Neote,K., DiGregorio,D., Mak,J.Y., Horuk,R. and Schall,T.J.
TITLE Molecular cloning, functional expression, and signaling
JOURNAL Characteristics of a C-C chemokine receptor
MEDLINE Cell 72 (3), 415-425 (1993)
FEATURES
    location/Qualifiers
    source
    1..1495
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /cell_line="differentiated HL60"
    1..1068
    /gene="CMKR-1"
    1..1068
    /gene="CMKR-1"
    /codon_start=1
    /product="C-C chemokine receptor type 1"
    /db_xref="PDB:9179985"
    /translation="METPNTEDYDTTTFDYGDATPCCKVNERAFGAQLPLPLYSLV

FVIGLVNILVVLVOYKRLKNNMTSIVLLNLAISDLLEFLTFPIFDWIDYKLKDDWVFG
DAMCKILSGEYVYTGLEYSEIFILLTIDRYLAIYHVAIPALPARYTFGVTSLIWAAL
AILASMPGLFLPQWETHHTCSLHPHESLREWKLFQALKNLGLVLPPLLMILIC
ITGILKILLRPNKSKKAVELIFVIMLIFLFWTYPYLNILISVQDFLFTHECQSQS
RHLDLAVQVTEVIAYTHCCVNPVYIAFVGERFRYKRLQLFHRVAVHLVLPFLSVLD
RLERSVSTSPSTGEHLSAGF"
BASE COUNT 348 a 389 c 361 g 397 t
ORIGIN
Query Match 5.5%; Score 61; DB 20; Length 1495;
Best Local Similarity 60.4%; Pred. No. 1.04e-27;
Matches 177; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
Db 121 TACTCCTTGGTATTGTTCATTTGGCCTGCTTGGAAACATCCTGGTGGTCTGCTGCTTGTG 180
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 TACTCCTTGGTATTGTTCATTTGGCCTGCTTGGGGAATATCTTGGTGGTATCACCTTTGCT 189
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CAATACAAGAGCTAAAAAATACATGACAGCATCTACCTCCCTGAACCTGGCCATTTCTGAC 240
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 TTTATAGAAGGCCAGGCTCTATGACAGACGCTCTATCTTGAACATGGCCATGACAGAC 249
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 CTGCTCTTCTTTCACGCTTCCCTTCTGATCGACTACAAGTTGAAGGATGACTGGGTT 300
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 ATCTCTTGTCTTACTCTCCCATTTCTGGCAGTGAGTGCATGCCCATGTTGCGTGGGTT 309
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TTTGGTGAATGATGTAAGATCTCTCTGGGTTTATTACACAGGCTTGTACAGCGAG 360
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 TTCACATGACGCTGCAAGTTGCTAAAGGCACTATGTCATCACTTAACTGCGGG 369
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 ATCTTTTTCATCCTCTGCTGACGATTGACAGGTAACCTGGCCATGCTCCAGC 413
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 ATGCTCTCTGACTGTCATTAGCATGACCGGTACATGCCCATTTGTACAGGC 422
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

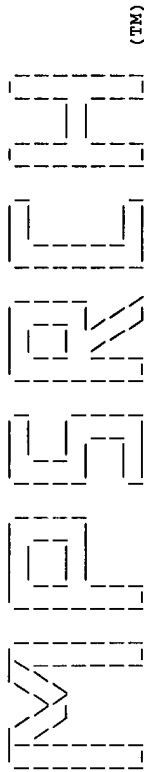
RESULT 14
LOCUS E13385 1609 bp DNA PAT 27-APR-1998
DEFINITION cDNA encoding human MIP-1 alpha /RANTES receptor.
ACCESSION E13385
NID 93252190
KEYWORDS JP 1997176048-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1609)
AUTHORS Honda,S. and Fujisawa,T.
TITLE PRODUCTION OF HUMAN MIP-1ALPHA/PANTES RECEPTOR PROTEIN AND USE
JOURNAL Patent: JP 1997176048-A 1 08-JUL-1997;
COMMENT TAKEDA CHEM IND LTD
OS Homo sapiens (human)
PN JP 1997176048-A/1
PD 08-JUL-1997
PF 28-DEC-1995 JP 1995342130
PI HONDA SUSUMU, FUJISAWA TOMOYUKI
PC A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,
PC A61K45/00,
PC A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,
PC A61K45/00,
PC A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,A61K45/00,
PC C07K14/705,
PC C12N5/10,C12N15/09,C12P21/02,C12Q1/00,(C12P21/02,C12R1/91), PC
(C12Q1/00,
PC C12R1/91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
Key Location/Qualifiers
FH
FT source
FT 1..1609
FT /organism="Homo sapiens"
FT /cell_type="leukemia"
FT /cell_line="U937"
```







\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Sat Nov 14 00:17:57 1998; MasPar time 167.79 Seconds  
907.323 Million cell updates/sec  
Tabular output not generated.

Title: >US-08-887-977-9  
Description: (1-1119) from US08887977.seq  
Perfect Score: 1119  
N.A. Sequence: 1 ARGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119  
Comp: TACAAAGCTGAGTGACACTT.....TCTTCGACTCAGAGGGATT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneseq32  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40

Statistics: Mean 9.061; Variance 4.959; scale 1.827

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	1118	99.9	1119	40	Human dendritic cell	0.00e+00
2	75	6.7	1900	11	Partial coding sequen	3.74e-34
3	75	6.7	2058	11	Putative seven trans	3.74e-34
4	75	6.7	2154	11	Epstein Barr virus in	3.74e-34
5	75	6.7	2160	11	Putative seven trans	3.74e-34
6	61	5.5	1085	33	Human MIP-1 alpha/RAN	1.99e-24
7	61	5.5	1495	11	C-C chemokine recepto	1.99e-24
8	61	5.5	2156	34	Human MIP-1 alpha/RAN	1.99e-24
9	62	5.5	2751	11	Seven transmembrane r	4.11e-25
10	57	5.1	1200	17	Recombinant high affi	1.05e-21
11	57	5.1	1200	5	Sequence encoding a h	1.05e-21
12	55	4.9	1071	36	Human CCR3 chemokine	2.33e-20
13	55	4.9	1116	23	CC-chemokine receptor	2.33e-20

14	55	4.9	1193	23	T31335	CC-chemokine receptor	2.33e-20
15	55	4.9	1689	32	T58783	Human C-C chemokine r	2.33e-20
16	55	4.9	1689	23	T31334	CC-chemokine receptor	2.33e-20
17	55	4.9	1915	34	T85162	Human chemokine recep	2.33e-20
18	55	4.9	5099	38	T93601	Human eosinophil eota	2.33e-20
19	51	4.6	1083	35	T96976	Human monocyte chemoa	1.09e-17
20	51	4.6	1979	15	O96298	Human monocyte chemoa	1.09e-17
21	51	4.6	2232	15	O96297	Human monocyte chemoa	1.09e-17
22	49	4.4	1176	17	O99950	Recombinant high affi	2.28e-16
23	49	4.4	1373	17	O99951	Recombinant high affi	2.28e-16
24	49	4.4	1607	23	T35277	Chemokine receptor K5	2.28e-16
25	49	4.4	1883	17	O99006	Interleukin 8 recepto	2.28e-16
26	49	4.4	1933	13	O80520	Interleukin-8 recepto	2.28e-16
27	49	4.4	1933	5	O29505	Interleukin-8 recepto	2.28e-16
28	48	4.3	1162	11	O66170	Seven transmembrane r	1.03e-15
29	48	4.3	2254	11	O66170	Seven transmembrane r	1.03e-15
30	46	4.1	1373	5	O30013	Sequence encoding a l	2.08e-14
31	45	4.0	1106	5	O30012	Sequence encoding a l	2.08e-14
32	45	4.0	1510	6	O38747	Human IL-8 receptor c	9.22e-14
33	45	4.0	1748	17	O99008	Interleukin 8 recepto	9.22e-14
34	44	3.9	1586	26	T44099	Human G-protein recep	4.06e-13
35	43	3.8	1106	17	O99952	Recombinant high affi	1.77e-12
36	43	3.8	1317	11	O66179	Seven transmembrane r	1.77e-12
37	43	3.8	1737	17	O99007	Chemokine superfamily	1.77e-12
38	43	3.8	1737	5	O29506	New platelet factor	1.77e-12
39	43	3.8	1737	13	O80521	Human monocyte PF4AR	1.77e-12
40	41	3.7	204	1	N81164	Base substituted E.co	3.31e-11
41	40	3.6	1244	7	O45656	Murine somatostatin r	1.41e-10
42	40	3.6	1299	11	O63919	Interleukin 8 (IL-8)	1.41e-10
43	39	3.5	1059	34	T85163	Macaque chemokine rec	5.95e-10
44	39	3.5	1080	21	T30385	Probe based on angiot	5.95e-10
45	38	3.4	514	23	T35278	Chemokine receptor cd	2.48e-09

ALIGNMENTS

RESULT 1  
ID VI5418 standard; cDNA; 1119 BP.  
AC VI5418;  
DT 11-JUN-1998 (first entry)  
DE Human dendritic cell chemokine receptor encoding cDNA.  
KW Human; thymus expressed chemokine; TECK; MIP-3alpha; MIP-3beta;  
KW receptor; dendritic cell; macrophage; inflammation; asthma; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..1098 /\*tag= a  
FT /\*product= "DC CR"  
FT /\*note= "dendritic cell chemokine receptor"  
FT /\*tag= b  
FT /\*note= "encodes His or Gln, but is stated as Gln in the protein (shown in W48086)."  
FT unsure 577..579  
PN WO9801557-A2.  
PD 15-JAN-1998.  
PR 02-JUL-1997; U10819.  
PR 04-JUN-1997; US-048593.  
PR 05-JUL-1996; US-675814.  
PR 11-OCT-1996; US-028329.  
PR (SCHE ) SCHERING CORP.  
PI Gish KC, Schall TJ, Vicari A, Wang W, Zlotnik A;  
DR WPI; 98-101054/09.  
DR P-PSDB: W48086.  
PT Novel chemokines, e.g. thymus expressed chemokine - used for  
PT treating inflammatory conditions including asthma.  
PS Claim 5; Page 92-94; 202pp; English.  
CC The present sequence encodes a human dendritic cell chemokine receptor.  
CC Antibodies which bind to the protein can be used in detecting or  
CC diagnosing various immunological conditions related to expression  
CC of the protein. The nucleic acid can be used for screening and  
CC isolating DNA clones for the chemokines, especially from other  
CC species. The chemokine can be used in the treatment of conditions  
CC associated with abnormal physiology or development, including



CC inflammatory conditions such as asthma.  
SQ Sequence 1119 BP; 261 A; 268 C; 257 G; 332 T;

Query Match 99.9%; Score 1118; DB 40; Length 1119;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1118; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgttttcgaactccagtgaaagtatttttggctcagtcagtcacatacttcatattactcagttg 60  
QY 1 ATGTTTTCGACTCCAGTGAAGATTATTTTGTCTAGTCAATATCTCATATTACTCAGTTG 60

Db 61 attctgagatttactgtctccttcagagaggtcagcagcttctccaggtcattttgtac 120  
QY 61 ATTCGAGATTTACTGTGCTTCCTTCAGAGAGGTTCAGCAGTCTCCAGGCTATTGTGTAC 120

Db 121 cgaattgctactccttgatctgttcttggccctccctggggaatatcttggtggtgac 180  
QY 121 CGAATTGCCACTCCTTGATCTGTGTTCTTGCCCTCCCTGGGGAATATCTGGTGTGATC 180

Db 181 accttgccttttataagaagccaggtctctatgacagacgtctctcttcttggaactggcc 240  
QY 181 ACCTTTGTCTTTTATAAGAGCCAGGCTCTATGACAGACGCTATCTCTGTAACATGGCC 240

Db 241 attgcagacatcctcttcttcttactctccattctcggcagtgagtcagtcactggt 300  
QY 241 ATTCGACATCTCTTGTGTTCTTACTCTCCATTCTGGGAGTGAGTCATGCCACTGGT 300

Db 301 gctgggttttcagcaatgcccagctgcaagttgctaaaggtcatctatgccaatcatt 360  
QY 301 GCGTGGGTTTCAGCAATGCCACGTGCAAGTGTCTAAAGGCATCTATGCCATCACTTT 360

Db 361 aactgctgggatgctcctcactgactgcatgcatgcatgcatgcatgcatgcatgcatg 420  
QY 361 AACTGCGGGATGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

Db 421 gcaactaagtcattccggctccgcatccagacacacacacacacacacacacacacac 480  
QY 421 GCGACTAAGTCTTCGGCTCCGATCCAGACACACTACCGCGCAGCAAAATCATCTGCTT 480

Db 481 gttgtgtgggggtcagtcacatcatctccagctcaactttgttctcaacaaataac 540  
QY 481 GTTGTGTGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

Db 541 aacacccaaagcagcagtgatgtgtgaaacccaaagtcacacacacacacacacacac 600  
QY 541 AACACCCAAAGCAGCAGTGTCTGTGAACCCAAAGTACCAAACTGTCTCGAGGCCATCAG 600

Db 601 tggagctgctgattgtggggttgagctactcttcttcttcttcttcttcttcttctt 660  
QY 601 TGGAGCTGCTGATGTTGGGGCTTGAGCTACTCTTGTGTTCTTTATCCCTTTGATGTT 660

Db 661 atgatatgtttacagcttcattgtcaaaacctgtgtgcaagctcagaattctctaaagg 720  
QY 661 ATGATATTTGTTACAGCTTCATTGTCTAAACCTTGTGTGCAAGCTCAGAACTCTAAAAG 720

Db 721 cacaagccatccgtgtaataatagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780  
QY 721 CACAAAGCCATCCGTGTAATCATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780

Db 781 cataacatgctcgtgtgtgacggtgcttaatttgggtaaaatgaacgactcctgcag 840  
QY 781 CATAACTGTCTCTGCTGTGTGACGGCTGCTAATTGGGTAATAAATGAACCGATCTTGC 840

Db 841 agcgaagcctaattgtctatcacgaaactgtcagagctcagagctcctgtcttctgcactgc 900  
QY 841 AGCGAAAGTAAATGGCTATACGAATACTGTACAGAAAGTCTTCCGTCTTCTGTCACATGC 900

Db 901 tgcctgaacctgtgctctacgcttttattggggagaaggttcagaataactttctgaag 960  
QY 901 TGCTGAACCTGTGCTCTACGCTTTTATTGGGAGAGAGTTTCAGAAACTACTTTCTGAAG 960

Db 961 atcttgaagacactgtgtgtgtgagaggaagtacaagtctcctcaggtctctctctgtcc 1020  
QY 961 ATCTTGAAGACCTGTGTTGTGTGAGAGGAAGTACAAGTCTCCTCAGGTTCTCTCTGTCC 1020

QY 961 ATCTTGAAGACCTGTGTTGTGTGAGAGGAAGTACAAGTCTCCTCAGGTTCTCTCTGTCC 1020

Db 1021 gggaggtactcagaaacattttctcggcagaccagtgagaccgagataaacgacaatgcy 1080  
QY 1021 GGGAGGTACTCAGAAACATTTCTCGCAGACCAGTGAGACCAGACGACATACGACATGCG 1080

Db 1081 tgcctcctcactatgtgatagaagctgagctcctcctaa 1119  
QY 1081 TCGTCTTCACTATGTGATAGAAGCTGAGTCTCCCTAA 1119

RESULT 2  
ID Q66162 standard; DNA; 1900 BP.  
AC Q66162;  
DT 02-FEB-1995 (first entry)  
DE Partial coding sequence of seven transmembrane receptor (V31).  
KW primer: seven transmembrane receptor; receptor; amplification; PCR;  
KW polymerase chain reaction; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT intron 1..168  
FT exon /\*tag- a  
FT exon 169..1245  
FT 3'utr /\*tag- b  
FT 1246..1900  
FT /\*tag- d  
PD WO9412635-A.  
PN 09-JUN-1994.  
PF 17-NOV-1993: U11153.  
PR 17-NOV-1992: US-977452.  
PA (ICOS-) ICOS CORP.  
PI Godiska R, Gray PW, Schweickart VL;  
DR WPI; 94-200264/24.  
PT DNA encoding seven transmembrane receptors - used to develop  
PT prods. for use as therapeutic or diagnostic agents for conditions  
PT involving the receptors.  
PS Example 3; Page 54-56; 100pp; English.  
CC Two primers (Q66148, Q66149) were used to amplify human genomic DNA  
CC purified from leukocytes. Approximately 1000 clones were isolated  
CC after the initial amplification reaction and probed with sequences  
CC specific for seven transmembrane receptors IL8R1, AT2R and R20.  
CC Clones which did not hybridise were then chosen for sequence  
CC analysis. Three new clones were identified that appeared to encode  
CC seven transmembrane receptor segments. Two more primers (Q66151,  
CC Q66152) were used to isolate a full length version of one of these  
CC clones designated V31 (See Q66153). This is the sequence of exon 3  
CC of the V31 genomic clone (along with partial intron sequences) and  
CC is the final exon of the clone. The TAG stop codon is found at  
CC positions 1243-1245 of the sequence.  
SQ Sequence 1900 BP; 431 A; 594 C; 458 G; 417 T;

Query Match 6.7%; Score 75; DB 11; Length 1900;  
Best Local Similarity 63.7%; Pred. No. 3.74e-34;  
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;

Db 248 tgtgtctcccaagagcagtcggaactt-taaagcctgtgtctcctccatcatgtactcc 306  
QY 76 TGTGTCTCTTGCAGGAGGTGAGGAGTCTTCCAGGCTATTGTGACGAATGTGCTACTCC 135

Db 307 atcaattgttttgggtcctactcgtggcaatgggtgtgtgtgtgtgtgtgtgtgtgtgt 366  
QY 136 TTGATCTGTGTTCTTGGCCCTCTCTGGGGAATATTCTGTGTGTGTGTGTGTGTGTGT 195

Db 367 aagagctcaagacacatgacacacacacacacacacacacacacacacacacacac 426  
QY 196 AAGAAGCCAGGCTATGACAGACGCTATCTCTTGAACATGGCCATTGACAGATCCTC 255

Db 427 ttctcctgaccccttcccttctgtggtcctacagcgcgcccaa--gtcc-tgggtcttcggt 483  
QY 256 TTGTGTTTACTCTCCATTCTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 315

Db 484 gtccacttttgaagctcatcttggccatctacaagatgagcttctcagtgagcagtc 543  
QY 484 GTCCACTTTTGAAGCTCATCTTGGCCATCTTGAAGATGAGCTTCTCAGTGTGAGCTC 543



QY 316 AATGCCACGTGCAAGTCTGCTAAAGSCATCTATGCGCATCAACTTTAACTCGGGATGCTG 375

Db 544 ctactctttgcatcagcattgacgctacgtggtggtccatgctccagc 590  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 376 CTCCTGACTTGCATTAGCATGACCGGTACATCGCCATTGTACAGGC 422

RESULT 3

ID Q66153 standard; DNA; 2058 BP.

AC Q66153; 1995 (first entry)

DE Putative seven transmembrane receptor (V31) coding sequence.

KW Primer: seven transmembrane receptor; receptor; amplification; PCR;

KW polymerase chain reaction; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT 166..11398

FT /\*tag= a

FT /product= V31 putative transmembrane receptor.

PN W09412635-A.

PD 09-JUN-1994.

PF 17-NOV-1993; U11153.

PR 17-NOV-1992; US-977452.

PA (ICOS-) ICOS CORP.

PI Godiska R, Gray PW, Schweickart VL;

DR WPI: 94-200264/24.

DR P-PSDB: R53743.

PT DNA encoding seven transmembrane receptors - used to develop

PT prods. for use as therapeutic or diagnostic agents for conditions

PT involving the receptors.

PS Example 2; Page 44-46; 100pp; English.

CC Two primers (Q66148, Q66149) were used to amplify human genomic DNA

CC purified from leukocytes. Approximately 1000 clones were isolated

CC after the initial amplification reaction and probed with sequences

CC specific for seven transmembrane receptors IL8R1, AT2R and R20.

CC Clones which did not hybridise were then chosen for sequence

CC analysis. Three new clones were identified that appeared to encode

CC seven transmembrane receptor segments. Two more primers (Q66151,

CC Q66152) were used to isolate a full length version of one of these

CC clones (this sequence) which was designated V31.

QY Sequence 2058 BP; 472 A; 632 C; 504 G; 448 T;

Query Match 6.7%; Score 75; DB 11; Length 2058;

Best Local Similarity 63.7%; Pred. No. 3.74e-34;

Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;

Db 401 tgtgtctccaagaagacgtgcggaactt-taaagcctgttctctccctatcatgtactcc 459  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 76 TGTGCTCTCTTGCGGCTCTCGGGATATTTCTGGTGATCACCTTTGCTTTTAT 195

Db 460 atcattgtttctgtggcctactgggcaatgggctgggtgtgtgacctaatactatttc 519  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 136 TTGATCTGTGTTTGGCTCTCGGGATATTTCTGGTGATCACCTTTGCTTTTAT 195

Db 520 aagaggtccaagaccataccataccctacccctgctcaacctggcgtggcagacatcttc 579  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 196 AAGAGGCCAGGTCTATGACACACGCTCTATCTCTTGACATGGCCATTGACAGATCCTC 255

Db 580 ttctcttgaccccttccctctggcctacagcgggccaa--gtcc-tgggtcttcggt 636  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 256 TTTGTCTTACTCTCCCATCTCGGGATATTTCTGGTGATCACCTTTGCTTTTAT 315

Db 637 gtccacttttgcaagctcatctttggcctacacagatgagcttcttcagtggtcagctc 696  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 316 AATGCCACGTGCAAGTCTGCTAAAGSCATCTATGCGCATCAACTTTAACTCGGGATGCTG 375

Db 697 ctactctttgcatcagcattgacgctacgtggtggtccatgctccagc 743  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 376 CTCCTGACTTGCATTAGCATGACCGGTACATCGCCATTGTACAGGC 422

RESULT 4

ID Q64125 standard; DNA; 2154 BP.

Q64125;

AC 03-FEB-1995 (first entry)

DE Epstein Barr virus induced (EBV-1) gene.

KW Epstein Barr virus; EBV; induction; detection; diagnosis;

KW lymphocytes; antigen; growth; differentiation; mediator;

KW infectious mononucleosis; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT 64..1200

FT /\*tag= a

FT /product= Epstein Barr virus induced polypeptide.

PN W09412519-A.

PD 09-JUN-1994.

PF 08-OCT-1993; U09636.

PR 25-NOV-1992; US-980518.

PA (BGHM) BRIGHAM & WOMENS HOSPITAL.

PI Birkenbach M, Kieff E;

DR WPI: 94-200183/24.

DR P-PSDB: R54079.

PT DNA coding for Epstein Barr Virus induced (EBV) polypeptide(s)

PT and antibodies to EBV, 2 and 3 - useful for detecting EBV by

PT hybridisation or by immunoassay

PS Claim 2; Page 52-54; 84pp; English.

CC EBV infected B lymphocytes recapitulate features of antigen

CC stimulation in enlarging, increasing RNA synthesis, expressing

CC activation antigens and adhesion molecules, secreting Ig and

CC proliferating. Unlike antigen stimulated B lymphocytes, EBV

CC infected B lymphocytes continue to proliferate (in vitro) as

CC immortalised lymphoblastoid cell lines. Because of the similar

CC effects of EBV and antigen, EBV induced genes are likely to include

CC mediators of antigen induced B lymphocyte growth or differentiation.

QY Sequence 2154 BP; 486 A; 647 C; 543 G; 478 T;

Query Match 6.7%; Score 75; DB 11; Length 2154;

Best Local Similarity 63.7%; Pred. No. 3.74e-34;

Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;

Db 203 tgtgtctccaagaagacgtgcggaactt-taaagcctgttctctccctatcatgtactcc 261  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 76 TGTGCTCTCTTGCGGAGGTGACGAGGTCTCTCAGGCTATTTGTACCGAATTCCTACTCC 135

Db 262 atcattgtttctgtggcctactgggcaatgggctgggtgtgtgacctaatactatttc 321  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 136 TTGATCTGTGTTTGGCTCTCGGGATATTTCTGGTGATCACCTTTGCTTTTAT 195

Db 322 aagaggtccaagaccataccataccctacccctgctcaacctggcgtggcagacatcttc 381  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 196 AAGAGGCCAGGTCTATGACACACGCTCTATCTCTTGACATGGCCATTGACAGATCCTC 255

Db 382 ttctcttgaccccttccctctggcctacagcgggccaa--gtcc-tgggtcttcggt 438  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 256 TTTGTCTTACTCTCCCATCTCGGGATATTTCTGGTGATCACCTTTGCTTTTAT 315

Db 439 gtccacttttgcaagctcatctttggcctacacagatgagcttcttcagtggtcagctc 498  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 316 AATGCCACGTGCAAGTCTGCTAAAGSCATCTATGCGCATCAACTTTAACTCGGGATGCTG 375

Db 499 ctactctttgcatcagcattgacgctacgtggtggtccatgctccagc 545  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 376 CTCCTGACTTGCATTAGCATGACCGGTACATCGCCATTGTACAGGC 422

RESULT 5

ID Q66160 standard; cDNA; 2160 BP.

AC Q66160;

DE Putative seven transmembrane receptor (V31-B) coding sequence (cDNA).

KW Primer: seven transmembrane receptor; receptor; amplification; PCR;

KW polymerase chain reaction; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT 64..1200

FT /\*tag= a







QY 130 TACTCCTTGTATCTGTTCTTGGCCCTCTGGGGAATATCTGGTGATACACCTTTGCT 189  
 Db 181 caatacaagagcgtataaaacatgacagcagctatcctcctgaacctggtcattttgac 240  
 QY 190 TTTTATAAGAGGCCAGGCTGTATGACAGACGCTTATCTCTTGAACATGGCCATTGCAGAC 249  
 Db 241 ctgctctctctgttcacgcttccctctctggtatcgactacaagtgaagatgactgggtt 300  
 QY 250 ATCCTCTTGTCTTACTCTCCCATCTCTGGGAGTGAGTCATGCCACTGGTGGTGGGTT 309  
 Db 301 ttgtgtgcatgcatgtgtaagatcctctctggtgtttattacacagggctgttacagcgag 360  
 QY 310 TTCAGCAATGCCAGCTGCAAGTGTCTAAAGGCATCTATGCCATCAACTTTAACTGCGGG 369  
 Db 361 atcttttcatcctcgtcgtgacgaggtgacaggtaccctggcctgctgcacgc 413  
 QY 370 ATGCTGCTCTGACTGTGATTAGCATGACATGGACCGGTACATGCCATGTACAGGC 422

## RESULT 8

ID T90384 standard; DNA; 2156 BP.  
 AC T90384;  
 DT 21-JAN-1998 (first entry)  
 DE Human MIP-1 alpha/RANTES receptor gene.  
 KW Macrophage inflammatory protein-1 alpha; MIP-1 alpha;  
 KW reduced upon activation normal T expressed and secreted; RANTES;  
 KW receptor; cytokine; antiinflammatory; inflammation; human; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 63..1130 /\*tag= a  
 FT FT  
 PN US5652133-A.  
 PD 29-JUL-1997.  
 PF 28-JAN-1993; 012988.  
 PR 28-JAN-1993; US-012988.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PI Murphy PW;  
 DR WPI; 97-392945/36.  
 DR P-PSDB; W26388.  
 PT MIP-1-alpha and RANTES receptor nucleic acid - used to develop  
 PT products for the detection of these cytokine(s) and their receptors,  
 PT particularly in inflammatory processes  
 PS Claim 1; Column 11-16; 12pp; English.  
 CC This nucleic acid sequence encodes a claimed receptor for macrophage  
 CC inflammatory protein-1 alpha (MIP-1 alpha) and regulated upon  
 CC activation normal T expressed and secreted (RANTES) protein. Also  
 CC claimed are: (1) a subsequence of the nucleic acid, having at least  
 CC 12 contiguous nucleotides; (2) a cell transformed or transfected  
 CC with the nucleic acid; and (3) purified MIP-1 alpha/RANTES receptor  
 CC polypeptide. The products can be used for detecting the MIP-1  
 CC alpha/RANTES receptor and polymorphisms in physiological samples.  
 CC In addition, the receptor can be expressed and used to assay for  
 CC MIP-1a/RANTES in biological samples. The quantitation of MIP-1  
 CC alpha/RANTES is useful for monitoring the levels of these cytokines  
 CC in a patient. Such measurements are useful in following the anti-  
 CC inflammatory effects of drugs and prospective usefulness of new anti-  
 CC inflammatory agents.  
 SQ Sequence 2156 BP; 538 A; 547 C; 504 G; 567 T;

Query Match 5.5%; Score 61; DB 34; Length 2156;  
 Best Local Similarity 60.4%; Pred. No. 1.99e-24;  
 Matches 177; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Db 183 tactccttggtattgttcattgctggtgttgaaacatcctggtgctggtcctggtctgtg 242  
 QY 130 TACTCCTTGTATCTGTTCTTGGCCCTCTGGGGAATATCTGGTGATACACCTTTGCT 189  
 Db 243 caatacaagagcgtataaaacatgacagcagctatcctcctgaacctggtcattttgac 302  
 QY 190 TTTTATAAGAGGCCAGGCTGTATGACAGACGCTTATCTCTTGAACATGGCCATTGCAGAC 249  
 Db 303 ctgctctctctgttcacgcttccctctctggtgtcgatcgactacaagtgaagatgactgggtt 362

QY 250 ATCCTCTTGTCTTACTCTCCCATCTCTGGGAGTGAGTCATGCCACTGGTGGTGGGTT 309  
 Db 363 ttgtgtgcatgcatgtgtaagatcctcctcctggtgtttattacacagggctgttacagcgag 422  
 QY 310 TTCAGCAATGCCAGCTGCAAGTGTCTAAAGGCATCTATGCCATCAACTTTAACTGCGGG 369  
 Db 423 atcttttcatcctcgtcgtgacgaggtgacaggtaccctggcctgctgcacgc 475  
 QY 370 ATGCTGCTCTGACTGTGATTAGCATGACCGGTACATGCCATGTACAGGC 422

## RESULT 9

ID Q66164 standard; DNA; 2751 BP.  
 AC Q66164;  
 DT 03-FEB-1995 (first entry)  
 DE Seven transmembrane receptor (V31) coding sequence.  
 KW Primer; seven transmembrane receptor; receptor; amplification; PCR;  
 KW polymerase chain reaction; ss.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT Intron 1..691 /\*tag= a  
 FT Exon 692..1771 /\*tag= b  
 FT poly\_a\_signal 2341..2348 /\*tag= c  
 PN W09412635-A.  
 PD 09-JUN-1994.  
 PF 17-NOV-1993; U11153.  
 PR 17-NOV-1992; US-977452.  
 PA (ICOS-) ICOS CORP.  
 PI Godiska R, Gray PW, Schweickart VL;  
 DR WPI; 94-200264/24.  
 DR P-PSDB; R53747.  
 PT DNA encoding seven transmembrane receptors - used to develop  
 PT prods. for use as therapeutic or diagnostic agents for conditions  
 PT involving the receptors.  
 PS Example 5; Page 58-60; 100pp; English.  
 CC Two primers (Q66148, Q66149) were used to amplify human genomic DNA  
 CC purified from leukocytes. Approximately 1000 clones were isolated  
 CC after the initial amplification reaction and probed with sequences  
 CC specific for seven transmembrane receptors IL8R1, AR2R and R20.  
 CC Clones which did not hybridise were then chosen for sequence  
 CC analysis. Three new clones were identified that appeared to encode  
 CC seven transmembrane receptor segments. Two more primers (Q66151,  
 CC Q66152) were used to isolate a full length version of one of the  
 CC human V31 clone (See Q66153). A fragment of the human clone was used  
 CC to isolate this V31 genomic clone of the mouse from a library.  
 SQ Sequence 2751 BP; 651 A; 747 C; 684 G; 669 T;

Query Match 5.5%; Score 62; DB 11; Length 2751;  
 Best Local Similarity 63.1%; Pred. No. 4.11e-25;  
 Matches 185; Conservative 0; Mismatches 105; Indels 3; Gaps 2;

Db 824 tatttctgcatcgtctgctggtcgccgctgctgcaacgggctggtgatactgaagcatcac 883  
 QY 130 TACTCCTTGTATCTGTTCTTGGCCCTCTGGGGAATATCTGGTGATACACCTTTGCT 189  
 Db 884 tattcaagaggctcaagaccatgacgggatacctgctcaacctggcctggcagac 943  
 QY 190 TTTTATAAGAGGCCAGGCTGTATGACAGACGCTTATCTCTTGAACATGGCCATTGCAGAC 249  
 Db 944 atccttttctcctaattctccttctccttctggtgctacagcagaagccaa--gtcc-tggatc 1000  
 QY 250 ATCCTCTTGTCTTACTCTCCCATCTCTGGGAGTGAGTCATGCCACTGGTGGTGGGTT 309  
 Db 1001 ttggtgctctaccctgtgtaaggcgcatctttggcatctataagttaagctcttcacggg 1060  
 QY 310 TTCAGCAATGCCAGCTGCAAGTGTCTAAAGGCATCTATGCCATCAACTTTAACTGCGGG 369  
 Db 1061 atgctgctgctctcatgcatcagcattgaccgctagcattgacccatgctccaggc 1113



QY 370 ATGCTGCTCTGACTTGTCATTAGCATGGACCGGTACATGCCATTGTACAGGC 422

RESULT 10  
ID Q99949 standard; DNA; 1200 BP.  
AC Q99949;  
DT 24-APR-1996 (first entry)  
DE Recombinant high affinity interleukin-8 receptor subtype A encoding DNA.  
KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;  
KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;  
KW systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;  
KW adult respiratory distress syndrome; neutrophil detection; ss.  
OS Oryctolagus cuniculus.

Key Location/Qualifiers  
FT 108..1175  
FT /\*tag= a  
FT /product= IL-8A\_receptor

PN W09525126-A1.  
PD 21-SEP-1995.  
PF 09-MAR-1995; U03032.  
PR 15-MAR-1994; US-210250.  
PR 02-MAY-1994; US-237937.  
PA (REPK ) REPLIGEN CORP.  
PA (UYBO-) UNIV BOSTON.  
PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;  
PI Witt DT;  
DR WPI; 95-336945/43.  
DR P-PSDB; R80950.  
PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -  
PT useful for treating inflammatory disorders, for detecting  
PT neutrophil(s) and for isolating IL-8 receptor from liq.mixt.  
PS Claim 2; Fig 1A-B; 74pp; English.  
CC Monoclonal antibodies were raised against recombinant interleukin-8  
CC (IL-8) receptor subtypes A and B from both human and rabbit sources  
CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8A) is  
CC a high affinity receptor and the B subtype receptor (IL-8B) is a  
CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to  
CC the IL-8 binding domain thus blocking its activation. The mAbs are  
CC useful for treating inflammatory disorders (see key words) and for  
CC detecting the presence of neutrophils in a biological sample. The  
CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.  
SQ Sequence 1200 BP; 234 A; 366 C; 295 G; 303 T;

Query Match 5.1%; Score 57; DB 17; Length 1200;  
Best Local Similarity 62.1%; Pred. No. 1.05e-21;  
Matches 185; Conservative 0; Mismatches 110; Indels 3; Gaps 3;

Db 257 ctatgcccctgtctctctgctgagcctgtggtggaacccctggtgatgctggtcactact 316  
QY 129 CTACTCTCTGTATCTGTGTTCTTGGCCCTCTGGGGAATATTTCTGGTGATCACCCTTGC 188  
Db 317 gtacagccggagcaacccgttcacccagcgtctacccgtgtaacccctgacatgcccga 376  
QY 189 TTTTATAGAGGCGAGCTCTATGACAGAGCTCTATCTCTTGAACATGCCATTGCAGA 248  
Db 377 cctgctcttttgcctgacatgcctctatctgggcccgtctc-caaggaaaaag-gc-tggat 433  
QY 249 CATCCTCTTTGTTCTTACTCTCCATCTCGGCGAGTGAGTCATGCCACTGTGCGTGGGT 308  
Db 434 ttccggcacgcccctgtgcaagggtgctcgtctgtggaagaaagcaactctacagtg 493  
QY 309 TTTTCAGCAATGCCAGCTGCAAGTGTCTAAAGGCATCTATGCCATCAACTTTAACTCGG 368  
Db 494 aatcctgctctggcctgacatgcctgagcctgacccctgacccctgacccctgacccctgact 551  
QY 369 GATGCTGCTCTGACTTGCTAGCATGAGACCGGTACATGCCATTGTACAGCGGACT 426

RESULT 11  
ID Q30011 standard; cDNA; 1200 BP.  
AC Q30011;  
DT 04-APR-1993 (first entry)  
DE Sequence encoding a high affinity recombinant rabbit interleukin-8

DE (IL-8) receptor polypeptide in F3R.  
KW IL-8 receptor polypeptide; G-protein-coupled receptor; ss.  
OS Oryctolagus cuniculus.  
Key Location/Qualifiers  
FT 108..1172  
FT /\*tag= a  
FT W09218641-A.  
PN 29-OCT-1992.  
PD 10-APR-1992; U02977.  
PF 10-APR-1991; US-685101.  
PR 09-JUL-1991; US-726606.  
PR 09-DEC-1991; US-803842.  
PA (REPK ) REPLIGEN CORP.  
PA (UYBO-) UNIV BOSTON.  
PI Navarro J, Thomas KM, Witt DP;  
PI WPI; 92-382123/46.  
DR P-PSDB; R28272.  
PT Recombinant mammalian interleukin-8 receptor - used for screening  
PT interleukin-8 binding antagonists, used to treat inflammation  
PS Disclosure; Fig 1; 71pp; English.  
CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit  
CC peritoneal neutrophils and used as a source of poly(A)+ RNA, to  
CC produce a rabbit neutrophil cDNA library. 250,000 recombinant  
CC plaques were screened for those which hybridized to an antisense  
CC oligonucleotide (Q30015). This probe was designed based on the  
CC sequence derived from the second transmembrane domain of G-protein-  
CC coupled receptors. After tertiary screening, six plaques were  
CC isolated. The insert of one of these plaques, termed F3R was of 2.5  
CC kb in size. This insert was sequenced. The protein deduced from  
CC the F3R clone demonstrates that it belongs to the family of  
CC G-protein-coupled receptors. The deduced protein sequence  
CC indicates seven putative transmembrane segments.  
SQ Sequence 1200 BP; 234 A; 366 C; 294 G; 304 T;

Query Match 5.1%; Score 57; DB 5; Length 1200;  
Best Local Similarity 62.1%; Pred. No. 1.05e-21;  
Matches 185; Conservative 0; Mismatches 110; Indels 3; Gaps 3;

Db 257 ctatgcccctgtctctctgctgagcctgtggtggaacccctggtgatgctggtcactact 316  
QY 129 CTACTCTCTGTATCTGTGTTCTTGGCCCTCTGGGGAATATTTCTGGTGATCACCCTTGC 188  
Db 317 gtacagccggagcaacccgttcacccagcgtctacccgtgtaacccctgacatgcccga 376  
QY 189 TTTTATAGAGGCGAGCTCTATGACAGAGCTCTATCTCTTGAACATGCCATTGCAGA 248  
Db 377 cctgctcttttgcctgacatgcctctatctgggcccgtctc-caaggaaaaag-gc-tggat 433  
QY 249 CATCCTCTTTGTTCTTACTCTCCATCTCGGCGAGTGAGTCATGCCACTGTGCGTGGGT 308  
Db 434 ttccggcacgcccctgtgcaagggtgctcgtctgtggaagaaagcaactctacagtg 493  
QY 309 TTTTCAGCAATGCCAGCTGCAAGTGTCTAAAGGCATCTATGCCATCAACTTTAACTCGG 368  
Db 494 aatcctgctctggcctgacatgcctgagcctgacccctgacccctgacccctgacccctgact 551  
QY 369 GATGCTGCTCTGACTTGCTAGCATGAGACCGGTACATGCCATTGTACAGCGGACT 426

RESULT 12  
ID T79096 standard; cDNA; 1071 BP.  
AC T79096;  
DT 13-MAR-1998 (first entry)  
DE Human CCR3 chemokine receptor coding sequence  
KW CCR3 chemokine; mouse; primer; PCR; amplification; antagonist; human;  
KW abnormal physiology; development; anti-viral; probe; hybridisation; ss.  
OS Homo sapiens.  
Key Location/Qualifiers  
FT 1..1071  
FT /\*tag= a  
FT /product= CCR3 chemokine receptor  
PN W09721812-A2.  
PD 19-JUN-1997.















\*\*\*\*\*

W P E R E H

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri Nov 13 23:16:17 1998; MasPar time 1258.39 Seconds  
Tabular output not generated.  
1364.368 Million cell updates/sec

Title: >US-08-887-977-9

Description: (1-1119) from US08887977.seq

Perfect Score: 1119 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCCCTAA 1119

N.A. Sequence: TACAAAGCTGAGTCACTT.....TCTTCGACTCAGAGGGATT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 1988617 seqs, 767163441 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb1-est55

Database: 1:em\_est10 2:em\_est11

Database: genbank-est107

3:gb\_est1 4:gb\_est10 5:gb\_est11 6:gb\_est12 7:gb\_est13

8:gb\_est14 9:gb\_est15 10:gb\_est16 11:gb\_est17

12:gb\_est18 13:gb\_est19 14:gb\_est20 15:gb\_est21

16:gb\_est21 17:gb\_est22 18:gb\_est23 19:gb\_est24

20:gb\_est23 21:gb\_est24 22:gb\_est25 23:gb\_est26 24:gb\_est7

25:gb\_est8 26:gb\_est9 27:gb\_gss

Statistics: Mean 11.175; Variance 1.882; scale 5.936

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
C 1	160	14.3	461	18	AI045155 UI-R-Cl-kk-c-08-0-UI.s	0.00e+00
C 2	133	11.9	312	13	AA790757 vvl8g09.r1 Soares mous	3.37e-239
C 3	121	10.8	492	16	AA889777 a150h01.s1 Soares_NFL	6.10e-211
C 4	80	7.1	635	22	W28677 50b1 Human retina CDNA	5.87e-117
C 5	66	5.9	252	13	AA754459 97SN1787 Rice Immature	2.62e-86
C 6	57	5.1	252	13	AA754459 97SN1787 Rice Immature	2.98e-67
C 7	52	4.6	247	13	AA754458 97SN1784 Rice Immature	5.89e-57
C 8	50	4.5	130	13	AA754458 97SN1784 Rice Immature	6.58e-53
C 9	50	4.5	247	13	AA754458 97SN1784 Rice Immature	6.58e-53
C 10	50	4.5	247	13	AA754458 97SN1784 Rice Immature	6.58e-53
C 11	42	3.8	261	4	AA331167 EST3521 Embryo, 8 wee	3.31e-37
C 12	42	3.8	316	4	AA296456 EST11038 Umbilical vei	3.31e-37
C 13	42	3.8	466	14	R15256 yf89d03.r1 Homo sapien	3.31e-37

14	38	3.4	289	20	T28268	EST34714 Homo sapiens	1.02e-29
15	38	3.4	391	26	AA153412	mq66a06.r1 Soares 2NM	1.02e-29
16	38	3.4	484	26	AA120018	mp93c11.r1 Soares 2NM	1.02e-29
17	38	3.4	639	13	AA816049	vi14h02.r1 Barstead mo	1.02e-29
18	38	3.4	793	25	AA182270	mt83f01.r1 Soares mous	1.02e-29
19	36	3.2	539	27	FR0019844	F.rubripes GSS sequenc	4.36e-26
20	36	3.2	798	17	AI007060	ua81c03.r1 Soares mous	4.36e-26
21	36	3.2	816	15	AA205847	zq50c02.r1 Stratagene	4.36e-26
22	35	3.1	249	20	H29103	ym31f07.r1 Homo sapien	2.65e-24
23	35	3.1	2275	11	AF034173	Homo sapiens ntcon2 co	2.65e-24
24	34	3.0	166	16	AA915469	vz31c03.r1 Soares 2NM	1.53e-22
25	34	3.0	342	10	AA621854	nq19g01.s1 NCI CGAP.Th	1.53e-22
26	34	3.0	382	14	R78657	vi74a05.r1 Homo sapien	1.53e-22
27	34	3.0	438	5	AA3866001	EST99759 Pancreas tumo	1.53e-22
28	34	3.0	2275	11	AF034173	Homo sapiens ntcon2 co	1.53e-22
29	33	2.9	145	14	R23114	yh27b12.r1 Homo sapien	8.36e-21
30	32	2.9	241	8	AA566795	ZF-A179 zebrafish adul	4.31e-19
31	32	2.9	360	27	FR0024014	F.rubripes GSS sequenc	4.31e-19
32	32	2.9	370	8	AA225739	nc17b08.s1 NCI CGAP.Pr	8.36e-21
33	32	2.9	521	8	AA479467	zvl1f10.r1 Soares NHM	4.31e-19
34	31	2.8	386	21	AA5296	yv16f09.s1 Homo sapien	2.08e-17
35	31	2.8	423	11	AA690396	vu52b02.r1 Soares mous	2.08e-17
36	31	2.8	424	5	AA035022	zk27c08.r1 Soares preg	2.08e-17
37	31	2.8	505	10	AA622439	nq40b08.s1 NCI CGAP.Co	2.08e-17
38	31	2.8	520	10	AA425767	z47g03.s1 Soares tota	2.08e-17
39	31	2.8	544	5	AA148399	z144c09.s1 Soares preg	2.08e-17
40	31	2.8	551	17	AA862993	og99c06.s1 NCI CGAP.Ki	2.08e-17
41	31	2.8	552	5	AA148398	z144c09.r1 Soares preg	2.08e-17
42	31	2.8	607	24	AA058871	z196h03.s1 Stratagene	2.08e-17
43	31	2.8	782	25	AA203362	zx54h11.r1 Soares feta	2.08e-17
44	30	2.7	410	6	C05136	Human Heart CDNA, clon	9.42e-16
45	30	2.7	478	26	AA189997	mt98e07.r1 Soares mous	9.42e-16

## ALIGNMENTS

RESULT 1	AI045155	461 bp	mRNA	EST	06-JUL-1998
LOCUS	UI-R-Cl-kk-c-08-0-UI.s1	UI-R-Cl	Rattus norvegicus	CDNA clone	
DEFINITION	UI-R-Cl-kk-c-08-0-UI 3', mRNA sequence.				
ACCESSION	AI045155				
NID	93291974				
KEYWORDS	EST.				
SOURCE	Norway rat.				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 461)				
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.				
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery				
JOURNAL	Genome Res. 6 (9), 791-806 (1996)				
MEDLINE	97044477				
COMMENT	Contact: Soares, MB Program for Rat Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: msoares@blue.weeg.uiowa.edu The sequence tag present in the CDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized adult Spleen library. CDNA Library Preparation: M. Fatma Bonaldo, Ph.D. Clone distribution: clones will be available through Research Genetics Seq primer: M13 Forward. Location/Qualifiers 1..461 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-Cl				



library is a subtracted library derived from the UI-R-C0 library, which is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-E1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C1) was constructed as follows: PCR amplified cDNA inserts from UI-R-C0 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C0 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C1 library. This procedure has been previously described (Bonaldio, Lennon and Soares, Genome Research 6: 791-806, 1996).

```

/db_xref="taxon:10116"
/clone="UI-R-C1-kl-c-08-0-UI"
/clone_lib="UI-R-C1"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
BASE COUNT 100 a 113 c 108 g 138 t 2 others
ORIGIN

```

```

Query Match 14.3%; Score 160; DB 18; Length 461;
Best Local Similarity 74.4%; Pred. NO. 0.00e+00;
Matches 264; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

```

```

Db 106 ACATGTAAGGAGGATCGGTGTCATTCTACAGTCTCAGTGGTCGCTGGAGATGT 165
Cp 1096 ACATAGTGAAGGAGGAGGATCGGTATATCGGGTCTCAGTGGTCGCGAGAAATGT 1037
Db 166 AGCTTTCTGAGTAAACCCGGGCACAGAAAGTAGGACCTTGCTCTCTCTCTCATAC 225
Cp 1036 ---TTTCTGAGTACCTCCGGCAGAGAGAGCTGAGGATTTGTACTTCTCTCACAC 980
Db 226 ACCACATCCTTCATGATCTTCATGAGTAGTCTCTGAATTTCTGTCCAAATGAAGCAT 285
Cp 979 ACCACAGGTCCTCAAGATCTTCAGAAAGTAGTTCTGAACCTCTGCCCAATAAAGCGT 920
Db 286 ACACACGGGTTGAGACAGAGTCGAGGAAGCCAGGACCTCAGCCACATCTCTGGCGT 345
Cp 919 AGAGCACAGGTTTCAGGCAGCAGTCGAGGAAGCCAGGACTTCTGTGACAGTTTTCTGAT 860
Db 346 ANGCGAGGGCTTCTCGCGCTGAGCTCGGGCCCATTTTCCCGTGTGTGCTGCAGTCA 405
Cp 859 AGCCATTAGCTTTCTGCTCTGCGAGATCGGTTCACTTTTACCACAAATAGACCGGCA 800
Db 406 CGAAGAGGACCATGTTGTGAGGATCTGACANGCCAGGACGAGAACACAGACAGCAGC 460
Cp 799 CAAGCAGGACCATGTTATGAGGAATCTGACAAGCCAGAAACACAGACACAGCAGC 745

```

```

RESULT 2
LOCUS 312 bp mRNA EST 06-FEB-1998
DEFINITION vw18g09.r1 Soares mouse mammary gland NBMGM Mus musculus cDNA clone
1244224 5', similar to SW:CKR6_HUMAN P51684 C-C CHEMOKINE RECEPTOR
TYPE 6 ;, mRNA sequence.
ACCESSION AA790757
NID 9280877
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryote: Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

## REFERENCE

1 (bases 1 to 312)  
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.  
 The WashU-HMI Mouse EST Project  
 Unpublished (1996)

## TITLE

CONTACT: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:657912

Possible reversed clone: similarity on wrong strand  
 Seq primer: -28ml3 rev2 ET from Amersham  
 High quality sequence stop: 217.

## FEATURES

```

Location/Qualifiers
1..312
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGGAATGGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M. Fatima Bonaldio."
/db_xref="taxon:10090"
/clone="1244224"
/clone_lib="Soares mouse mammary gland NBMGM"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
BASE COUNT 70 a 80 c 78 g 84 t
ORIGIN

```

```

Query Match 11.9%; Score 133; DB 13; Length 312;
Best Local Similarity 71.4%; Pred. NO. 3.37e-239;
Matches 222; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```

```

Db 1 ACCAAGTCTTTCCGGTAGCTCCAGAACACTGACGACAGTAAGTCTATCTGTATGGCA 60
Qy 424 ACTAGTCAATCCGGTCCGATCCAGAACACTACCGCGAGCAAAATCACTGCTTGT 483
Db 61 GTGTGTTTCATCTCCATCATCTCAAGCCCTTACATTTAGCTTCAACAGAGATACGAG 120
Qy 484 GTGTGGGGCTGTAGTCACTATCTCCAGCTCAACTTTTGTCTTCAACCAAAATACAAC 543
Db 121 CTGCAGGATCGTGATGTCTGTGAGCCAGGTACAGTCTGTCTCAGAGCCCATCAGTGG 180
Qy 544 ACCCAAGGAGGATGTCTGTGAACCCCAAGTCTCGGAGCCCATCAGTGG 603
Db 181 AGGCTGTGGTATGGAGCTGGAGCTGTTCTTTGGTCTTTCACCCCTTTGCTGTATTATG 240
Qy 604 AGCTGCTGATTTGGGGCTTGAGTACTCTTTGGTTCTTTATCCCTTGTGATGTTCAAG 663
Db 241 GTGTTCTGCTATCTGTCAATTAATCAGAACTTGGTGCAGGCCCCAGAACTTCCAGGAC 300
Qy 664 ATATTGTTTACACGTTTCATTGTCAAAACCTTGGTGTCAAGCTCAGAAATCTTAAAGGCAC 723
Db 301 AGAACATCCG 311
| | | | |

```



```

Oy 724 AAAGCCATCCG 734

RESULT 3
LOCUS AA889777 492 bp mRNA EST 03-APR-1998
DEFINITION al50h01.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1460785 3' similar to SW:CKR6_HUMAN P51684 C-C CHEMOKINE
RECEPTOR TYPE 6 ; mRNA sequence.
ACCESSION AA889777
NID 93016656
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 492)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncilogap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -40m3 fwd. Et from Amersham
High quality sequence stop: 109.
FEATURES
Location/Qualifiers
source
1..492
/organism="Homo sapiens"
/note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19w, testis NHT, and B-cell
NCL-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Donaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1460785"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
BASE COUNT 134 a 119 c 117 g 122 t
ORIGIN
Query Match 10.8%; Score 121; DB 16; Length 492;
Best Local Similarity 80.8%; Pred. No. 6.10e-211;
Matches 172; Conservative 0; Mismatches 39; Indels 2; Gaps 2;
Db 282 TTAGGGGGCCCCACCTTATCAATGTGA-GGACGGCGCATGTCGTAACTCGGGT 340
Cp 1119 TTAGGGAGACTCAGCTTCTATCATAGTGAAGGACGACGCTATGCTGTCGGGT 1060
Db 341 CTCATGTGCTCCGAGAAAGGTTTCGGGGGAACTCGGGGAAGGGAAGCCTGAGGA 400
Cp 1059 CTCATGTGCTCGGAGAAATGTTTCTGAGTACCTCCCGGCACAGGAGGCTGAGGA 1000
Db 401 CTGGTACTTCTCCAGAGCACCAAGGTC-TTCAAGATCTTCCGAAAGTAGTTTCGAA 459
Cp 999 CTGTGACTTCTCTCACACACACAGGTCCTTCAAGATCTTCAGAAAGTAGTTTCGAA 940
Db 460 CTCTGCGCAATAAAGGGTTAAGCACAGGGTT 492
Cp 939 CTCTGCGCAATAAAGCGTAGAGCACAGGTT 907

RESULT 4
LOCUS W28677 635 bp mRNA EST 08-MAY-1996
DEFINITION 50b1 Human retina cDNA randomly primed sublibrary Homo sapiens
cDNA, mRNA sequence.
ACCESSION W28677
NID gl308625
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 635)
AUTHORS Macke, J., Smallwood, P. and Nathans, J.
TITLE Adult Human Retina cDNA
JOURNAL Unpublished (1996)
COMMENT Contact: Dr. Jeremy Nathans
Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics
Johns Hopkins School of Medicine
725 North Wolfe Street, Baltimore, MD 21205
Tel: 410 953 4678
Fax: 410 614 0827
Email: jeremy_nathans@mail.bs.jhu.edu
Clones from this library are NOT available.
PCR Primers
FORWARD: CTTTTCAGCAAGTTCAGCTCGTTAAGT
BACKWARD: GAGTGGCTTATGAGTATTTCTTCAGGGTAA
Seq primer: GGGTAAAGCAAGAAAGATT.
FEATURES
Location/Qualifiers
source
1..635
/organism="Homo sapiens"
/note="Organ: eye; Vector: lambda gt10; Site_1: EcoRI;
Site_2: EcoRI; the library used for sequencing was a
sublibrary derived from a human retina cDNA library.
Inserts from retina cDNA library DNA were isolated,
randomly primed, PCR amplified, size-selected, and cloned
into lambda gt10. Individual plaques were arrayed and
used as templates for PCR amplification, and these PCR
products were used for sequencing."
/db_xref="taxon:9606"
/clone_lib="Human retina cDNA randomly primed sublibrary"
/sex="mixed (males and females)"
/tissue_type="retina"
/dev_stage="adult"
/lab_host="E. coli strain K802"
BASE COUNT 134 a 136 c 143 g 128 t 94 others
ORIGIN
Query Match 7.1%; Score 80; DB 22; Length 635;
Best Local Similarity 71.4%; Pred. No. 5.87e-117;
Matches 177; Conservative 0; Mismatches 58; Indels 13; Gaps 13;
Db 209 TTTCTNTTCATGCGGAGGANGAGGNATNN-CGATATCTGNGGCTCAGTGGTCT-CCG 266
Cp 1104 TTTCTATCATAGTAGGAGGACGACGATTCCTGTTATCTCGGCTCAGTGGTCTGCC 1045
Db 267 ANATATCTNN-CTCAGTACC-CCCGGAACAGAGAGANNNGANN-CTNGNTCT-CANNCT 322
Cp 1044 AGAATGTTTCTGAGTACCTCCCGGCACAGAGAGACCTGAGGACTGTACTTCTCT 985
Db 323 NANACACACAGGTCCT-CAAGATCT-CAGAATGAGTGTCTTNN-CTNTCT-CCCA-TCAG 377
Cp 984 CACACACACAGGTCCTCAAGATCTTCAGAAAGTCTTCTGAACTTCTGCCCAATAA 925
Db 378 ANNGGGGTGCACAGNCTNAGNAGAGCAGACACN-CCAGNACN-CTNGACANTTNC 435
Cp 924 AGCTAGACGACGGGTTCAGGCAGCAGTGCAGGAAAGCCAGGACTTCTGTGACAGTTT 865
Db 436 CGTNTGCG 443
Cp 864 CGTATAGC 857

```



```

5
RESULT 5
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cdna Library Oryza sativa
CDNA clone 97SN1787, mRNA sequence.
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 252)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT
Contact: Eun M.Y.
Department of CytoGenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@s20.osti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
Location/Qualifiers
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cdna library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cdna Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOL8"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 5.9%; Score 66; DB 13; Length 252;
Best Local Similarity 10.3%; Pred. No. 2.62e-86;
Matches 24; Conservative 120; Mismatches 89; Indels 1; Gaps 1;
Db 18 WMTSTYBCHGNBWWVCVASHGNYMNVNCTBRGTHCDCKNNV-WSTMWGTWNNBVNSGD 76
QY 453 ACTACCGCGCAGCAAAATCATCTGCTTGTGTGGGGGCTGCAGTCATCATCTCCAG 512
Db 77 WYWBVBNTKVDVGNHTRFCSWRBVRMAHYHDYNCBBYNNNDYHMHBBYBGTGCT 136
QY 513 CTCAACTTTTGTCTTCAACCAAAATCAACACCAAGGAGCGAGCTGCTGTAACCCAA 572
Db 137 CTMWCWBHNTKCTASGWHSTNTYDKSSTNTWGTBTSYDKSMHGYWCS-BBVKYHTKVST 196
QY 573 GTACCAAACTGCTCGGAGCCATCAGTGGAGCTGCTGATGTTGGGCTTGAGCTACT 632
Db 197 TRATRYTCVRKYCVMMTKKVKYHVBBGCHBTDSCKTKMTMMNKHVMTST 250
QY 633 CTTTGGTTTCTTATCCCTTGATGTTTCATGATATTTTGTACACGTTTCATGT 686

6
RESULT 6
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cdna Library Oryza sativa
CDNA clone 97SN1787, mRNA sequence.
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 247)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

```

```

rice.
SOURCE Oryza sativa
ORGANISM Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 252)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT
Contact: Eun M.Y.
Department of CytoGenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@s20.osti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
Location/Qualifiers
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cdna library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cdna Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOL8"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 5.1%; Score 57; DB 13; Length 252;
Best Local Similarity 11.6%; Pred. No. 2.98e-67;
Matches 27; Conservative 117; Mismatches 86; Indels 3; Gaps 3;
Db 21 TSYBCHGNBWWVCVASHGNYMNVNCTBRGTHCDCKNNVWSTMWGTWNNBVNSGDWHY 79
Cp 463 TCGCGGTAGTGTCTTGATCGGAGCGGAGTACTAGTCCCTGATCAATGGCATGT 404
Db 80 WEVBNTKVDVGNHTRFCSWRBVRMAHYHDYNCBBYNNNDYHMHBBYBGTGCTM 139
Cp 403 ACCGTCCTCATGTAATGCAAGTCAGGAGCATCCCGCAGTAAAGTTGATGGCATAGA 344
Db 140 WCWBHNTKCTASGWHSTNTYDKSSTNTWGTBTSYDKSMHGYWCS-BBVKYHTKVSTTR 198
Cp 343 TGCCTTTAGCAACTGTCAGCTGGCATGCTGCTGAAACCCACCGACGATGGCATCA 284
Db 199 ATRYTCVRKYCVMMTKKVKYHVBBGCHBTDSCKTKMTMMNKHVMTST 251
Cp 283 CTGCCCAAGTGGGAGAGTAAGAACAAAGAG-GATGCTCTGCAATGCCCATGTT 232

7
RESULT 7
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cdna Library Oryza sativa
CDNA clone 97SN1784, mRNA sequence.
ACCESSION AA754458
NID 92801164
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 247)
REFERENCE
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

```







Waterston, R.  
The WashU-HHMI Mouse EST Project  
Unpublished (1996)

Contact: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:657910

Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand  
Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 1.  
Location/Qualifiers  
1. .130  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/note="Organ: mammary gland; Vector: pT7T3D-Pac  
(Pharmacia) with a modified polylinker; Site\_1: Not I;  
Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGAGCGCGCGGAATGGTGTGTTTTTTTTTTTTTTTT  
T 3'] double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT7T3 vector.  
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M. Fatima  
Bonaldo."  
/db\_xref="taxon:10090"  
/clone="1244222"  
/clone\_lib="Soares mouse mammary gland NBMXG"  
/sex="male"  
/tissue\_type="mammary gland"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"

BASE COUNT 35 a 35 c 27 g 33 t  
ORIGIN

Query Match 4.5%; Score 50; DB 13; Length 130;  
Best Local Similarity 73.9%; Pred. No. 6.58e-53;  
Matches 85; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

Db 2 AAATCTTCGGTGCAGTCTCCAGAACTGACGACAGTAAAGTCACTGTGTGCGCATG 61  
QY 427 AAGTCAATCCGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTTGTG 486

Db 62 TGGTTCATCTCCATCATCATCTCAAGCCCTACATTTAGTCTTCAACAGAAATAC 116  
QY 487 TGGGGCTGTCAGTCATCTCAGCTCAACTTTT-GTCTTCAACCAAAATAC 540

RESULT 10  
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998  
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII CDNA Library Oryza sativa  
ACCESSION AA754458  
NID 92801164  
KEYWORDS EST.  
SOURCE rice.  
ORGANISM Oryza sativa  
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;  
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;  
Poales; Poaceae; Oryza.  
1 (bases 1 to 247)  
REFERENCE  
AUTHORS Nahm, B.H., Kim, J.K., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P.,  
Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y.,  
Lee, M.C. and Eun, M.Y.  
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

Unpublished (1998)  
Contact: Eun M.Y.  
Department of Cytogenetics  
National Inst. of Agri. Sci. and Tech, RDA  
Suwon, Kyunggi-do, Korea  
Tel: 82 331 290 0301  
Fax: 82 331 290 0307  
Email: myeunesun20.asti.re.kr  
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji  
University, Yongin, Korea. 449-728 bhnahm@bloserver.myongji.ac.kr  
Seq primer: M13 Reverse Primer.  
Location/Qualifiers  
1. .247  
/organism="Oryza sativa"  
/cultivar="Milyang23"  
/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:  
XhoI; Directional cDNA library inserted into lambda ZAPII  
vector at 5' end with EcoRI and 3' end with Xho I site."  
/db\_xref="taxon:4530"  
/clone="97SN1784"  
/clone\_lib="Rice Immature Seed Lambda ZAPII CDNA Library"  
/tissue\_type="Immature Seed"  
/dev\_stage="5 days after pollination"  
/lab\_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others  
ORIGIN

Query Match 4.5%; Score 50; DB 13; Length 247;  
Best Local Similarity 13.0%; Pred. No. 6.58e-53;  
Matches 31; Conservative 108; Mismatches 97; Indels 2; Gaps 2;

Db 1 HWDCTMTVWRCGCCBAAWNNKHTHMTBBWCVRVGTGTTNKGKNGRTTWNDCSDNA 60  
Cp 670 AAAATATCATGACATCAAGGATAAAGAACCAAGTAGCTCAAGCCCCCAATCA 611

Db 61 HCR-YTVBMYARSKYGYGTBYYSNNVNTGTTGGTGYKTTVNVHSGNWNCSNVVYVWB 119  
Cp 610 GCAGCTTCACCTGATGGCTCGACACAGTTTGTGTACITGGTTTCACAGACATCGCTGC 551

Db 120 TAYCDBYBHDRAHVDDTRCTNDRGYCNVTASDNGTSATKRVGTDKTSDCGGGGWRK 179  
Cp 550 CTGGGTGTGTATTTTGTGTTGAAGAAAAGTTGAGCTGGA-GATGATGACTGACAGC 492

Db 180 VYGSBBYBRGVNVVRYTTSMTDKSTKMBSDMRSSRVHYGRWBNKRGMSRNW 237  
Cp 491 CCCACACAAAGGAGATGATTTTGTGCGCGGTAGTGTCTGTGATCGGACCGGA 434

RESULT 11  
LOCUS AA331167 261 bp mRNA EST 21-APR-1997  
DEFINITION ES135211 Embryo, 8 week 1 Homo sapiens CDNA 5' end similar to  
vasoactive intestinal peptide receptor RDC1, mRNA sequence.  
ACCESSION AA331167  
NID 91983628  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Hominidae;  
Homo.  
REFERENCE 1 (bases 1 to 261)  
AUTHORS Adams, M.D., Kervage, A.R., Fleischmann, R.D., Fuldner, R.A.,  
Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,  
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C.,  
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,  
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,  
Glodek, A., Gnehm, C.L., Hanha, M.C., Hedblom, E., Hinkle, P.S., Jr.,  
Kelley, J.M., Kelley, J.C., Liu, L.-I., Marnaros, S.M., Merrick, J.M.,  
Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,  
Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,  
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,  
Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,







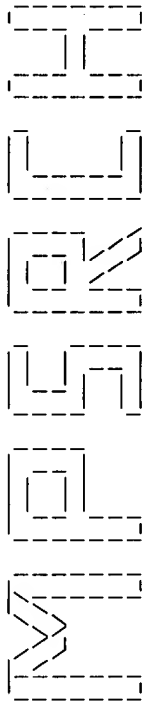








\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Sat Nov 14 00:36:39 1998; MasPar time 44.20 Seconds  
Tabular output not generated. 1181.065 Million cell updates/sec

Title: >US-08-887-977-9  
Description: (1-1119) from US08887977.seq

Perfect Score: 1119 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119  
N.A. Sequence: Comp: TACAAAGCTGAGGTCACCTT.....TCTTCGACTCAGAGGATT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 8.680; Variance 4.376; scale 1.984

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	75	6.7	1900	2	PCT-US93-1 Sequence 18, Applicati	2.18e-38
2	75	6.7	2058	2	PCT-US93-1 Sequence 6, Applicatio	2.18e-38
3	75	6.7	2154	2	PCT-US93-0 Sequence 1, Applicatio	2.18e-38
4	75	6.7	2160	2	PCT-US93-1 Sequence 14, Applicati	2.18e-38
5	61	5.5	2156	1	US-08-012- Sequence 1, Applicatio	9.02e-28
6	62	5.5	2751	2	PCT-US93-1 Sequence 23, Applicati	1.61e-28
7	58	5.2	7218	1	US-08-232- Sequence 14, Applicati	1.54e-25
8	57	5.1	1200	2	PCT-US93-0 Sequence 1, Applicatio	8.47e-25
9	53	4.7	1200	2	PCT-US92-0 Sequence 1, Applicatio	7.35e-22
10	51	4.6	1979	2	PCT-US95-0 Sequence 3, Applicatio	2.10e-20
11	51	4.6	2232	2	PCT-US95-0 Sequence 1, Applicatio	2.10e-20
12	49	4.4	1176	2	PCT-US93-0 Sequence 2, Applicatio	5.83e-19
13	49	4.4	1373	2	PCT-US92-0 Sequence 3, Applicatio	5.83e-19
14	49	4.4	1373	2	PCT-US93-0 Sequence 3, Applicatio	5.83e-19
15	49	4.4	1883	1	US-08-202- Sequence 1, Applicatio	5.83e-19
16	49	4.4	1933	1	US-08-418- Sequence 1, Applicatio	5.83e-19
17	49	4.4	1933	1	US-08-410- Sequence 1, Applicatio	5.83e-19
18	49	4.4	1933	1	US-08-076- Sequence 1, Applicatio	5.83e-19
19	49	4.4	1933	2	PCT-US94-0 Sequence 1, Applicatio	5.83e-19
20	49	4.4	1933	1	US-08-410- Sequence 1, Applicatio	5.83e-19

21	48	4.3	1161	2	PCT-US93-1 Sequence 31, Applicati	3.05e-18
22	48	4.3	2254	2	PCT-US93-1 Sequence 27, Applicati	3.05e-18
23	45	4.0	1510	1	US-07-759- Sequence 4, Applicatio	4.16e-16
24	45	4.0	1748	1	US-08-202- Sequence 8, Applicatio	4.16e-16
25	43	3.8	1106	2	PCT-US92-0 Sequence 5, Applicatio	1.06e-14
26	43	3.8	1106	2	PCT-US95-0 Sequence 4, Applicatio	1.06e-14
27	43	3.8	1317	2	PCT-US93-1 Sequence 45, Applicati	1.06e-14
28	43	3.8	1737	1	US-08-076- Sequence 3, Applicatio	1.06e-14
29	43	3.8	1737	1	US-08-202- Sequence 4, Applicatio	1.06e-14
30	43	3.8	1737	2	PCT-US94-0 Sequence 2, Applicatio	1.06e-14
31	40	3.6	652	2	PCT-US92-0 Sequence 2, Applicatio	1.28e-12
32	40	3.6	1244	1	US-07-816- Sequence 7, Applicatio	1.28e-12
33	39	3.5	1080	2	PCT-US95-1 Sequence 3, Applicatio	6.23e-12
34	36	3.2	1572	1	US-08-041- Sequence 5, Applicatio	6.88e-10
35	36	3.2	1572	1	US-08-417- Sequence 5, Applicatio	6.88e-10
36	36	3.2	1679	1	US-08-202- Sequence 6, Applicatio	6.88e-10
37	36	3.2	1679	2	PCT-US94-0 Sequence 3, Applicatio	6.88e-10
38	36	3.2	1679	1	US-08-076- Sequence 5, Applicatio	6.88e-10
39	32	2.9	1351	1	US-07-816- Sequence 5, Applicatio	2.86e-07
40	29	2.6	69	2	PCT-US93-1 Sequence 1, Applicatio	2.30e-05
41	29	2.6	215	1	US-08-238- Sequence 5, Applicatio	2.30e-05
42	29	2.6	720	2	PCT-US93-1 Sequence 35, Applicati	2.30e-05
43	28	2.5	74	2	PCT-US95-1 Sequence 100, Applicati	9.59e-05
44	28	2.5	105	1	US-07-865- Sequence 13, Applicati	9.59e-05
45	28	2.5	1842	1	US-08-446- Sequence 1, Applicatio	9.59e-05

## ALIGNMENTS

RESULT 1  
ID PCT-US93-11153-18 STANDARD; DNA; UNC; 1900 BP.  
AC xxxxxx  
DT  
Sequence 18, Application PC/TUS9311153  
Sequence 18, Application PC/TUS9311153  
GENERAL INFORMATION:  
CC APPLICANT: Godiska, Ronald  
CC APPLICANT: Gray, Patrick W.  
CC APPLICANT: Schweikart, Vicki L.  
CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors  
CC NUMBER OF SEQUENCES: 64  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
CC ADDRESSEE: Bicknell  
CC STREET: 6300 Sears Tower, 233 South Wacker Drive  
CC CITY: Chicago  
CC STATE: Illinois  
CC COUNTRY: USA  
CC ZIP: 60606  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/11153  
CC FILING DATE:  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 07/977,452  
CC FILING DATE: 17-NOV-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Noland, Greta E.  
CC REGISTRATION NUMBER: 35,302  
CC REFERENCE/DOCKET NUMBER: 31794  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (312) 474-6300  
CC TELEFAX: (312) 474-0448  
CC TELE: 25-3856  
CC INFORMATION FOR SEQ ID NO: 18:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1900 base pairs  
CC TYPE: nucleic acid











CC ADDRESSEE: Townsend and Townsend and Hourie and Crew  
CC STREET: One Market Plaza, Steuart Tower, Suite 2000  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94610  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/012,988A  
CC FILING DATE: 19930128  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Weber, Kenneth A.  
CC REGISTRATION NUMBER: 31,677  
CC REFERENCE/DOCKET NUMBER: 15280-118  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 415-543-9600  
CC TELEFAX: 415-543-5043  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2156 base pairs  
CC TYPE: NUCLEIC ACID  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC FEATURE:  
CC NAME/KEY: primer\_bind  
CC LOCATION: 259..275  
CC  
CC NAME/KEY: primer\_bind  
CC LOCATION: complement (868..884)  
CC  
CC NAME/KEY: CDS  
CC LOCATION: 63..1128  
CC SEQUENCE 2156 BP; 538 A; 547 C; 504 G; 567 T; 0 OTHER.

Query Match 5.5%; Score 61; DB 1; Length 2156;  
Best Local Similarity 60.4%; Pred.No. 9,02e-28;  
Matches 177; Conservative 0; Mismatches 116; Indels 0; Gaps 0;  
Db 183 TACTCCTGGTATTTGTCATTTGGCTGGTGGTGAACATCTCTGGTGGTCTGGTCTTTGTG 242  
QY 130 TACTCCTGGTATCTGTTCTTGGCTCTCTGGGGAATATCTTGGTGGTATCACCTTTGCT 189  
Db 243 CAATACAGAGGCTAAACACATGACAGCATCTACCTCTGAACCTGGCCATTTCTGAC 302  
QY 190 TTTTATAAGAGGCGCAGGTCTATGACAGACGCTCTATCTTTGAACATGGCCATTCGAC 249  
Db 303 CTGCTCTCTGTTACGCTTCCCTCTGGATCGACTACAAAGTTGAAGATGACTGGGTT 362  
QY 250 ATCCTCTTTGTTTCTTACTCTCCCACTTGGCGAGTGTATGCGACCTGGTGGGTT 309  
Db 363 TTTGGGTATGCCATGTTAAGATCTCTCTGGGTTTATTACAGAGGCTTGTACAGCGAG 422  
QY 310 TTCAGCAATGCCAGTGCGAAGTTGCTAAAGGCATCTATGCCATCAACTTAACTGCGGG 369  
Db 423 ATCTTTTCAATCTCTGCTGACGATGACAGGATGACCTGGCCATCGTCCAGC 475  
QY 370 ATGCTGCTCTGACTTGCATTAGCATGACGCGGTATACGCGCATTTGTACAGGC 422

## RESULT 6

ID PCT-US93-11153-23 STANDARD; DNA; UNC; 2751 BP.

AC xxxxxx

DT

DE Sequence 23, Application PC/TUS9311153

CC Sequence 23, Application PC/TUS9311153

CC GENERAL INFORMATION:

CC APPLICANT: Godiska, Ronald

CC APPLICANT: Gray, Patrick W.  
CC APPLICANT: Schweikart, Vicki L.  
CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors  
CC NUMBER OF SEQUENCES: 64  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
CC ADDRESSEE: Bicknell,  
CC STREET: 6300 Sears Tower, 233 South Wacker Drive  
CC CITY: Chicago  
CC STATE: Illinois  
CC COUNTRY: USA  
CC ZIP: 60606  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/11153  
CC FILING DATE:  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 07/977,452  
CC FILING DATE: 17-NOV-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Noland, Greta E.  
CC REGISTRATION NUMBER: 35,302  
CC REFERENCE/DOCKET NUMBER: 31794  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (312) 474-6300  
CC TELEFAX: (312) 474-0448  
CC TELEX: 25-3856  
CC INFORMATION FOR SEQ ID NO: 23:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2751 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC FEATURE:  
CC NAME/KEY: intron  
CC LOCATION: 1..691  
CC  
CC NAME/KEY: exon  
CC LOCATION: 692..1771  
CC  
CC NAME/KEY: CDS  
CC LOCATION: 692..1768  
CC  
CC NAME/KEY: polyA\_signal  
CC LOCATION: 2341..2348  
CC  
CC SEQUENCE 2751 BP; 651 A; 747 C; 684 G; 669 T; 0 OTHER.

Query Match 5.5%; Score 62; DB 2; Length 2751;  
Best Local Similarity 63.1%; Pred.No. 1.61e-28;  
Matches 185; Conservative 0; Mismatches 105; Indels 3; Gaps 2;

Db 824 TATTCTGTCATCTGCTCGTGGCCCTGCTCGCAACGGGCTGGTGATCTAGCTACATC 883  
QY 130 TACTCCTGATCTGTTCTTGGCCCTCTGGGGAATATTCTGGTGGTATCACCTTTGCT 189  
Db 884 TATTTCAAGAGGCTCAAGACCATGACGGATACCTACCTGCTCAACCTGGCCGTCGACAG 943  
QY 190 TTTTATAAGAGGCGCAGGTCTATGACAGAGCTATCTCTTGACATGGCCATTCGACAG 249  
Db 944 ATCCTTTTCTCTTAATTTCTTCCCTTTCTGGGCTTACAGGCAAGCCAA--GTCC-TGATC 1000  
QY 250 ATCCTCTTTGTTTCTTACTCTCCCATCTGGGAGTGTAGTATGCCACTGGTGGGTT 309  
Db 1001 TTTGGCGTCTACCTGTGTAAAGGCATCTTTGGCATCTATAAGTTAGCTTCTTCACGGG 1060  
QY 310 TTCAGCAATGCCAGCTGCAAGTTGCTAAAGGCATCTATGCCATCAACTTAACTGCGGG 369











```
CC LOCATION: 81..1160
SQ SEQUENCE 1979 BP; 530 A; 435 C; 451 G; 563 T; 0 OTHER.

Query Match 4.6%; Score 51; DB 2; Length 1979;
Best Local Similarity 61.2%; Pred. No. 2.10e-20;
Matches 180; Conservative 0; Mismatches 111; Indels 3; Gaps 2;

Db 224 CTACTCGCTGGTGTTCATCTTTGGTGGGCAACATGCTGGTGGTCTCATCTTAAT 283
QY 129 CTACTCGCTGGTGTTCATCTTTGGTGGGCAACATGCTGGTGGTCTCATCTTAAT 188
Db 284 AAACGTGCAAAAGCTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 343
QY 189 TTTTATGAAGAGCCAGCTGATGACAGAGCTGCTCTTGAACATGSCCATTCGAGA 248
Db 344 TCTGCTTTTCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
QY 249 CATCTCTTTGTTTACTCTCCCATCTGGCAGTGGTGGTGGTGGTGGTGGTGGTGG 308
Db 401 CTTTGGGAATGCAATGTCGAAATTTACAGGGCTGATACATCGGTTATTTGGCGG 460
QY 309 TTTGAGCAATGCGAGTGAAGTTGCTTAAAGCATCTATGCGCATCACTTTAACTG 368
Db 461 AATCTCTTCTCATCTCTCTGACATGATAGATACCTGGCTATTGTCCATGC 514
QY 369 GATGCTGCTCTGACTTGCTGATGATGATGATGATGATGATGATGATGATGATG 422

RESULT 11
ID PCT-US95-00476-1 STANDARD; DNA; UNC; 2232 BP.
AC xxxxxx
DT
DE Sequence 1, Application PC/TUS9500476
CC Sequence 1, Application PC/TUS9500476
CC GENERAL INFORMATION:
CC APPLICANT: The Regents of the University of California
CC TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
CC TITLE OF INVENTION: PROTEIN RECEPTORS
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Robbins, Berliner & Carson
CC STREET: 201 N. Figueroa Street, 5th Floor
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90012-2628
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/00476
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Berliner, Robert
CC REGISTRATION NUMBER: 20,121
CC REFERENCE/DOCKET NUMBER: 5555-291
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 310-977-1001
CC TELEFAX: 310-977-1003
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2232 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FEATURE:
```

```
CC NAME/KEY: CDS
CC LOCATION: 40..1161
SQ SEQUENCE 2232 BP; 602 A; 464 C; 508 G; 658 T; 0 OTHER.

Query Match 4.6%; Score 51; DB 2; Length 2232;
Best Local Similarity 61.2%; Pred. No. 2.10e-20;
Matches 180; Conservative 0; Mismatches 111; Indels 3; Gaps 2;

Db 183 CTACTCGCTGGTGTTCATCTTTGGTGGGCAACATGCTGGTGGTCTCATCTTAAT 242
QY 129 CTACTCGCTGGTGTTCATCTTTGGTGGGCAACATGCTGGTGGTCTCATCTTAAT 188
Db 243 AAACGTGCAAAAGCTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 302
QY 189 TTTTATGAAGAGCCAGGCTGATGACAGAGCTGCTCTTGAACATGSCCATTCGAGA 248
Db 303 TCTGCTTTTCTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359
QY 249 CATCTCTTTGTTTACTCTCCCATCTGGCAGTGGTGGTGGTGGTGGTGGTGGTGG 308
Db 360 CTTTGGGAATGCAATGTCGAAATTTACAGGGCTGATACATCGGTTATTTGGCGG 419
QY 309 TTTGAGCAATGCGAGTGAAGTTGCTTAAAGCATCTATGCGCATCACTTTAACTG 368
Db 420 AATCTCTTCTCATCTCTCTGACATGATAGATACCTGGCTATTGTCCATGC 473
QY 369 GATGCTGCTCTGACTTGCTGATGATGATGATGATGATGATGATGATGATGATG 422

RESULT 12
ID PCT-US95-03032-2 STANDARD; DNA; UNC; 1176 BP.
AC xxxxxx
DT
DE Sequence 2, Application PC/TUS9503032
CC Sequence 2, Application PC/TUS9503032
CC GENERAL INFORMATION:
CC APPLICANT: Repligen Corporation
CC APPLICANT: the Trustees of Boston University
CC TITLE OF INVENTION: ANTIBODIES TO INTERLEUKIN-8 RECEPTORS AND
CC TITLE OF INVENTION: METHODS OF USE
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC COMPUTER: IBM PS/2 Model 502 or 55SX
CC OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
CC SOFTWARE: WordPerfect (Version 5.0)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/03032
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Berliner, Robert
CC REGISTRATION NUMBER: 20,121
CC REFERENCE/DOCKET NUMBER: 5555-291
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 310-977-1001
CC TELEFAX: 310-977-1003
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2232 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FEATURE:
```











\*\*\*\*\*

WIREH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Nov 14 00:39:25 1998; MasPar time 901.46 Seconds  
Tabular output not generated. 1314.048 Million cell updates/sec

Title: >US-08-887-977-9  
Description: (1-1119) from US08887977.seq

Perfect Score: 1119  
N.A. Sequence: 1 ATGTTTCGACTCCAGTGAA.....AGAAAGCTGAGTCTCCCTAA 1119  
Comp: TACAAAGCTGAGTCACTT.....TCTTCGACTCAGAGGATT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 1639711 seqs, 529296662 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-pending  
1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005  
8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82  
15:U83 16:U84A 17:U84B 18:U85 19:U86 20:U87 21:U88  
22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU7  
29:NEWU8 30:NEWU9

Statistics: Mean 10.319; Variance 3.340; scale 3.090

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1119	100.0	1119	19	US-08-675-Sequence 7, Applicatio	0.00e+00
2	1119	100.0	1119	21	US-08-887-Sequence 9, Applicatio	0.00e+00
3	1104	98.7	1255	23	US-09-023-Sequence 993, Applicat	0.00e+00
4	1104	98.7	1679	23	US-09-016-Sequence 1097, Applicat	0.00e+00
5	1070	95.6	1137	23	US-09-023-Sequence 970, Applicat	0.00e+00
6	246	22.0	375	20	US-08-706-Sequence 3273, Applicat	4.84e-26
C 7	96	8.6	578	19	US-08-624-Sequence 14, Applicati	2.29e-80
C 8	96	8.6	578	18	US-08-569-Sequence 14, Applicati	2.29e-80
C 9	76	6.8	383	16	US-08-446-Sequence 8, Applicatio	1.19e-56
C 10	76	6.8	383	16	US-08-446-Sequence 8, Applicatio	1.19e-56
C 11	76	6.8	383	16	US-08-446-Sequence 8, Applicatio	1.19e-56
C 12	76	6.8	383	15	US-08-311-Sequence 8, Applicatio	1.19e-56
C 13	76	6.8	383	17	US-08-451-Sequence 8, Applicatio	1.19e-56
C 14	76	6.8	383	16	US-08-446-Sequence 8, Applicatio	1.19e-56
15	75	6.7	1500	23	US-09-023-Sequence 1465, Applicat	1.74e-55
16	75	6.7	1900	14	US-08-245-Sequence 18, Applicati	1.74e-55

17	75	6.7	2058	11	US-07-977- Sequence 6, Applicatio	1.74e-55
18	75	6.7	2058	14	US-08-245- Sequence 6, Applicatio	1.74e-55
19	75	6.7	2154	11	US-07-980- Sequence 1, Applicatio	1.74e-55
20	75	6.7	2154	11	US-07-980- Sequence 1, Applicatio	1.74e-55
21	75	6.7	2154	11	US-07-980- Sequence 1, Applicatio	1.74e-55
22	75	6.7	2154	15	US-08-352- Sequence 1, Applicatio	1.74e-55
23	75	6.7	2154	11	US-07-980- Sequence 1, Applicatio	1.74e-55
24	75	6.7	2154	23	US-09-023- Sequence 1094, Applicatio	1.74e-55
25	75	6.7	2154	23	US-09-016- Sequence 1214, Applicatio	1.74e-55
26	75	6.7	2160	14	US-08-245- Sequence 14, Applicatio	1.74e-55
27	73	6.5	578	18	US-08-569- Sequence 14, Applicatio	3.66e-53
28	73	6.5	578	19	US-08-624- Sequence 9, Applicatio	3.66e-53
C	29	65	340	19	US-08-617- Sequence 14, Applicatio	5.61e-44
30	65	5.8	383	16	US-08-446- Sequence 8, Applicatio	5.61e-44
31	65	5.8	383	15	US-08-311- Sequence 8, Applicatio	5.61e-44
32	65	5.8	383	16	US-08-446- Sequence 8, Applicatio	5.61e-44
33	65	5.8	383	17	US-08-451- Sequence 8, Applicatio	5.61e-44
34	65	5.8	383	16	US-08-446- Sequence 8, Applicatio	5.61e-44
35	65	5.8	383	16	US-08-446- Sequence 8, Applicatio	5.61e-44
36	64	5.7	2577	23	US-09-016- Sequence 1095, Applicatio	7.66e-43
C	37	63	339	19	US-08-617- Sequence 8, Applicatio	1.04e-41
38	61	5.5	1094	22	US-08-902- Sequence 1, Applicatio	1.86e-39
39	61	5.5	1495	23	US-09-023- Sequence 1021, Applicatio	1.86e-39
40	61	5.5	1495	1	PCT-US93-1 Sequence 8, Applicatio	1.86e-39
41	61	5.5	1495	15	US-08-308- Sequence 8, Applicatio	1.86e-39
42	61	5.5	1495	23	US-09-016- Sequence 1190, Applicatio	1.86e-39
43	62	5.5	2085	14	US-08-245- Sequence 65, Applicatio	1.39e-40
44	61	5.5	2156	23	US-09-023- Sequence 1247, Applicatio	1.86e-39
45	62	5.5	2751	14	US-08-245- Sequence 23, Applicatio	1.39e-40

ALIGNMENTS

RESULT 1  
ID US-08-675-814-7 STANDARD; DNA; UNC; 1119 BP.  
AC xxxxxx  
DT  
DE Sequence 7, Application US/08675814  
CC Sequence 7, Application US/08675814  
CC GENERAL INFORMATION:  
CC APPLICANT: Gish, Kurt C.  
CC APPLICANT: Schall, Thomas J.  
CC APPLICANT: Vicari, Alain  
CC APPLICANT: Zlotnik, Albert  
CC TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS  
CC NUMBER OF SEQUENCES: 14  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: DNAX Research Institute  
CC STREET: 901 California Avenue  
CC CITY: Palo Alto  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94304-1104  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA: US/08/675.814  
CC APPLICATION NUMBER: 05-JUL-1996  
CC FILING DATE: 05-JUL-1996  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Ching, Edwin P.  
CC REGISTRATION NUMBER: 34,090  
CC REFERENCE/DOCKET NUMBER: DX0589  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 415-852-9196  
CC TELEFAX: 415-496-1200  
CC INFORMATION FOR SEQ ID NO: 7:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1119 base pairs  
CC TYPE: nucleic acid



CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: CDNA  
CC FEATURE:

CC NAME/KEY: CDS

CC LOCATION: 1..1098

SQ SEQUENCE 1119 BP; 262 A; 268 C; 257 G; 332 T; 0 OTHER.

Query Match 100.0%; Score 1119; DB 19; Length 1119;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ATGTTTCGACTCCAGTGAAGATATTTTGTGTCAGTCAATCTCATATTACTCAGTTG 60  
QY 1 ATGTTTCGACTCCAGTGAAGATATTTTGTGTCAGTCAATCTCATATTACTCAGTTG 60  
Db 61 ATTCTGAGATGTTACTGTCTCTTGCAGGAGGTCAGGAGTCTCTCCAGGCTATTGTTAC 120  
QY 61 ATTCTGAGATGTTACTGTCTCTTGCAGGAGGTCAGGAGTCTCTCCAGGCTATTGTTAC 120  
Db 121 CGAATTCCTTACTCCTTGATCTGTGTTCTTGGCCCTCTCGGGAATATTCTGGTGTGATC 180  
QY 121 CGAATTCCTTACTCCTTGATCTGTGTTCTTGGCCCTCTCGGGAATATTCTGGTGTGATC 180  
Db 181 ACCTTTCCTTTTATAAGAGGCCAGGCTCTATGACAGAGCTCTATCTCTTGAACATGSCC 240  
QY 181 ACCTTTCCTTTTATAAGAGGCCAGGCTCTATGACAGAGCTCTATCTCTTGAACATGSCC 240  
Db 241 ATTGCAGACATCCTTTTGTCTTCTTACCTCTCCCATCTGGCAGTGAGTCATGCCACTGGT 300  
QY 241 ATTGCAGACATCCTTTTGTCTTCTTACCTCTCCCATCTGGCAGTGAGTCATGCCACTGGT 300  
Db 301 GCGTGGGTTTTCAGCAATGCCAGTCAAGTGTGTAAAGCATCTATGCCATCAACTTT 360  
QY 301 GCGTGGGTTTTCAGCAATGCCAGTCAAGTGTGTAAAGCATCTATGCCATCAACTTT 360  
Db 361 AACTGGGGATGTCCTGCTGACTTGTATGATGATGACGACGCTATCGCCATGTTACAG 420  
QY 361 AACTGGGGATGTCCTGCTGACTTGTATGATGATGACGACGCTATCGCCATGTTACAG 420  
Db 421 GCGACTAAGTCATTCGGCTCCGATCCAGACACTACCGCGCAGCAAAATCATCTGCTT 480  
QY 421 GCGACTAAGTCATTCGGCTCCGATCCAGACACTACCGCGCAGCAAAATCATCTGCTT 480  
Db 481 GTTGTGTGGGGGTGTGTCAGTCATCTCCAGCTCAACTTTTGTCTTCAACCAAAATAC 540  
QY 481 GTTGTGTGGGGGTGTGTCAGTCATCTCCAGCTCAACTTTTGTCTTCAACCAAAATAC 540  
Db 541 AACACCCAGGACGATGTCGTGAACCCAGTACCAAACTGCTCGGAGCCCATCAGG 600  
QY 541 AACACCCAGGACGATGTCGTGAACCCAGTACCAAACTGCTCGGAGCCCATCAGG 600  
Db 601 TGAAGTCTGCTGATGTTGGGGCTTGAGCTACTCTTGTGTTTCTTATCCCTTGTATGTC 660  
QY 601 TGAAGTCTGCTGATGTTGGGGCTTGAGCTACTCTTGTGTTTCTTATCCCTTGTATGTC 660  
Db 661 ATGATATTTTGTACAGTTCAATTTGTCAGAACTTGTGCAAGCTCAGAAATCTAAAAGG 720  
QY 661 ATGATATTTTGTACAGTTCAATTTGTCAGAACTTGTGCAAGCTCAGAAATCTAAAAGG 720  
Db 721 CACAAAGCCATCCGTTAATCATAGCTGTGGTGTGTTGTTTCTGGCTGTGAGATTCTT 780  
QY 721 CACAAAGCCATCCGTTAATCATAGCTGTGGTGTGTTGTTTCTGGCTGTGAGATTCTT 780  
Db 781 CATACATGCTGCTGTGAGGCTGCTAAATTTGGTAAATGAACCGATCTGCCAG 840  
QY 781 CATACATGCTGCTGTGAGGCTGCTAAATTTGGTAAATGAACCGATCTGCCAG 840  
Db 841 AGCGAAAGCTAATTTGGCTATACGAAACTGTCTACAGAGCTCTGGCTTCTCGTCACTGC 900  
QY 841 AGCGAAAGCTAATTTGGCTATACGAAACTGTCTACAGAGCTCTGGCTTCTCGTCACTGC 900  
Db 901 TGCCTGAACCTGTGCTCTACGCTTTTATTTGGGCGAGAAGTTTCAGAAACTACTTTCTGAAG 960

QY 901 TGCCTGAACCTGTGCTCTACGCTTTTATTTGGCAGAGTTTCAGAAACTACTTTCTGAAG 960  
Db 961 ATCTTGAAGGACCTGTGTGTGTGAGAAGAGTACAAAGTCTCAGGCTTCTCCTGTGCC 1020  
QY 961 ATCTTGAAGGACCTGTGTGTGTGAGAAGAGTACAAAGTCTCAGGCTTCTCCTGTGCC 1020  
Db 1021 GGGAGTACTCAGAAACATTTCTCGGAGACCCAGTGAACCCGAGATACGACATGCG 1080  
QY 1021 GGGAGTACTCAGAAACATTTCTCGGAGACCCAGTGAACCCGAGATACGACATGCG 1080  
Db 1081 TCGTCTCTCACATGTGATAGAAAGCTGAGTCTCCCTAA 1119  
QY 1081 TCGTCTCTCACATGTGATAGAAAGCTGAGTCTCCCTAA 1119

RESULT 2

ID US-08-887-977-9 STANDARD; DNA; UNC; 1119 BP.

AC xxxxx

DT

DE Sequence 9, Application US/08887977

CC Sequence 9, Application US/08887977

CC GENERAL INFORMATION:

CC APPLICANT: Wang, Wei

CC APPLICANT: Gish, Kurt C.

CC APPLICANT: Schall, Thomas J.

CC APPLICANT: Vicari, Alain P.

CC APPLICANT: Zlotnik, Albert

CC TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS

CC NUMBER OF SEQUENCES: 19

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: DNAX Research Institute

CC STREET: 901 California Avenue

CC CITY: Palo Alto

CC STATE: California

CC COUNTRY: USA

CC ZIP: 94304-1104

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent In Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC FILING DATE: 03-JUL-1997

CC CLASSIFICATION:

CC APPLICATION NUMBER: US 60/021,644

CC FILING DATE: 05-JUL-1996

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 60/028,329

CC FILING DATE: 11-OCT-1996

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Ching, Edwin P.

CC REGISTRATION NUMBER: 34,090

CC REFERENCE/DOCKET NUMBER: DX0589K1

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 650-852-9192

CC TELEFAX: 650-496-1200

CC INFORMATION FOR SEQ ID NO: 9:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1119 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: cdna

CC FEATURE:

CC NAME/KEY: CDS

CC LOCATION: 1..1095

SQ SEQUENCE 1119 BP; 262 A; 268 C; 257 G; 332 T; 0 OTHER.

Query Match 100.0%; Score 1119; DB 21; Length 1119;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;



	Matches	1119;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Db	1	ATGTTTTCGACTCCAGCTGAAGATTATTTTGTGTCAGTCAATACCTTCATATTACTCAGTTG	60							
QY	1	ATGTTTTCGACTCCAGCTGAAGATTATTTTGTGTCAGTCAATACCTTCATATTACTCAGTTG	60							
Db	61	ATTCTGAGATGTTACTGTGTCCTTCGAGGAGTTCAGGCAGTTCCTCCAGCTATTGTGTAC	120							
QY	61	ATTCTGAGATGTTACTGTGTCCTTCGAGGAGTTCAGGCAGTTCCTCCAGCTATTGTGTAC	120							
Db	121	CGAATTGCCCTACTCTTGATCTGTGTTCTTGGCTCCTGGGAATATTCCTGGTGGTGATC	180							
QY	121	CGAATTGCCCTACTCTTGATCTGTGTTCTTGGCTCCTGGGAATATTCCTGGTGGTGATC	180							
Db	181	ACCTTTGCTTTTTATAGAAGGCCAGGCTCTATGACAGAGCTCTATCTTTGAACATGGCC	240							
QY	181	ACCTTTGCTTTTTATAGAAGGCCAGGCTCTATGACAGAGCTCTATCTTTGAACATGGCC	240							
Db	241	ATTGCAGACATCCTCTTTGTTCTTACTCTCCCATTTCTGGSCAGTGAATGCCACTGGT	300							
QY	241	ATTGCAGACATCCTCTTTGTTCTTACTCTCCCATTTCTGGSCAGTGAATGCCACTGGT	300							
Db	301	GCCTGGGTTTTACGAATGCCAGTCCAGCTGCAAGTCTGCTAAAAGGCATCTATGCCATCAACTTT	360							
QY	301	GCCTGGGTTTTACGAATGCCAGTCCAGTGCAGAGTTCCTAAAAGGCATCTATGCCATCAACTTT	360							
Db	361	AACTTGGGGATGCTGCTCTGACTTTCATTAAGCATGACCGGTACATCGCCATTGTACAG	420							
QY	361	AACTTGGGGATGCTGCTCTGACTTTCATTAAGCATGACCGGTACATCGCCATTGTACAG	420							
Db	421	GCAGCTAAGTCAATTCCGGCTCCGATCCAGAACTACTACCGCGCAGCAAAATCACTGCCCTT	480							
QY	421	GCAGCTAAGTCAATTCCGGCTCCGATCCAGAACTACTACCGCGCAGCAAAATCACTGCCCTT	480							
Db	481	GTGTGTGGGGCTGTGAGTCATCATCTCCAGCTCAACTTTTCTCTACCAAAATATAC	540							
QY	481	GTGTGTGGGGCTGTGAGTCATCATCTCCAGCTCAACTTTTCTCTACCAAAATATAC	540							
Db	541	AACACCCAAGCGCAGGATGCTGTGAACCAAGTACCAACTGTCTCGGAGCCCATCAGG	600							
QY	541	AACACCCAAGCGCAGGATGCTGTGAACCAAGTACCAACTGTCTCGGAGCCCATCAGG	600							
Db	601	TGGAAGCTGCTGATGTTTGGGGCTTGAGCTACTCTTTGGTTCTTTATCCCTTTGATGTTTC	660							
QY	601	TGGAAGCTGCTGATGTTTGGGGCTTGAGCTACTCTTTGGTTCTTTATCCCTTTGATGTTTC	660							
Db	661	ATGATATTTTGTACAGTTCATTGTCAAAACCTTGGTCAAGCTCAGAAATCTTAAAGG	720							
QY	661	ATGATATTTTGTACAGTTCATTGTCAAAACCTTGGTCAAGCTCAGAAATCTTAAAGG	720							
Db	721	CACAAAGCCATCCGCTGAATCATAGCTGTGGTGCTGTGTTTCTGGCTGTGTGATTCCT	780							
QY	721	CACAAAGCCATCCGCTGAATCATAGCTGTGGTGCTGTGTTTCTGGCTGTGTGATTCCT	780							
Db	781	CATAAATGTCCTGCTGTTGTGACGGCTGCTGAATTTGGTAAAAATGAACCGATCCTGCCAG	840							
QY	781	CATAAATGTCCTGCTGTTGTGACGGCTGCTGAATTTGGTAAAAATGAACCGATCCTGCCAG	840							
Db	841	AGCGAAAAGCTAATTGGCTATACGAAACTGTACAGAGTCTCGCTTTCCTGGCACTGC	900							
QY	841	AGCGAAAAGCTAATTGGCTATACGAAACTGTACAGAGTCTCGCTTTCCTGGCACTGC	900							
Db	901	TGCTGAACCTGTGCTCTACGCTTTTATTGGGCAGAGTTTCAGAAACTACTTTCTGGAAG	960							
QY	901	TGCTGAACCTGTGCTCTACGCTTTTATTGGGCAGAGTTTCAGAAACTACTTTCTGGAAG	960							
Db	961	ATCTTGAAGGACCTGTGGTGTGTGAGAAGGATACAAAGTCTCTAGGCTTCTCCTGTGCC	1020							
QY	961	ATCTTGAAGGACCTGTGGTGTGTGAGAAGGATACAAAGTCTCTAGGCTTCTCCTGTGCC	1020							
Db	1021	GGGAGGTACTCAGAAACATTTCTCGGCAGACCAAGTGTAGACCCAGATACAGCAATGGG	1080							
QY	1021	GGGAGGTACTCAGAAACATTTCTCGGCAGACCAAGTGTAGACCCAGATACAGCAATGGG	1080							

```

Db 1081 TCGTCCTTCTACTGTGATAGAGAAGCTGAGTCTCCCTAA 1119
      |||||||
QY 1081 TCGTCCTTCTACTGTGATAGAGAAGCTGAGTCTCCCTAA 1119
      |||||||

RESULT 3
ID US-09-023-655-993 STANDARD; DNA; UNC; 1255 BP.
AC xxxxxx
AD DT
DE Sequence 993, Application US/09023655
CC Sequence 993, Application US/09023655
CC GENERAL INFORMATION:
CC APPLICANT: Cocks, Benjamin G.
CC APPLICANT: Susan G. Stuart
CC APPLICANT: Jeffrey J. Seilhamer
CC TITLE OF INVENTION: COMPOSITION FOR THE DETECTION O
CC TITLE OF INVENTION: EXPRESSION
CC NUMBER OF SEQUENCES: 1508
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/023,655
CC FILING DATE: HERewith
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Zeller, Karen J.
CC REGISTRATION NUMBER: 37,071
CC REFERENCE/DOCKET NUMBER: PA-0001 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 855-0555
CC TELEFAX: (650) 845-4166
CC INFORMATION FOR SEQ ID NO: 993:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1255 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: GENBANK
CC CLONE: g1668737
CC SO SEQUENCE 1255 BP; 302 A; 294 C; 284 G; 375 T; 0 OTHER.

```

Query Match	98.7%	Score 1104;	DB 23;	Length 1255;
Best Local Similarity	99.6%	pred. No. 0.00e+00;		
Matches 1114;	Conservative 0;	Mismatches 4;	Indels 1;	Gaps 1;
Db	89	ATGTTTTGCACTCCAGTGAAGATTATTTGTGTCAGTCAATACTTCATATTACTCAGTTG	148	
Qy	1	ATGTTTTGCACTCCAGTGAAGATTATTTGTGTCAGTCAATACTTCATATTACTCAGTTG	60	
Db	149	ATTCTGAGATGTTACTGTGTCCTTCGACAGAGTTCAGGCAGTTCTCCAGGCTATTTGTAC	208	
Qy	61	ATTCTGAGATGTTACTGTGTCCTTCGACAGAGTTCAGGCAGTTCTCCAGGCTATTTGTAC	120	
Db	209	CGA-TTGGCTACTCCTTGATCTGTGTCCTTTGGCCTCCTGGGGAATATTCCTGGTGGTGATC	267	
Qy	121	CGAATTGCCCTACTCCTTGATCTGTGTCCTTTGGCCTCCTGGGGAATATTCCTGGTGGTGATC	180	







Db 675 GCGACTAAGTTCATCCGGCTCCGATCCAGAACACTACCGCGAGCAAAATCATCTGCCTT 734  
QY 421 GCGACTAAGTTCATCCGGCTCCGATCCAGAACACTACCGCGAGCAAAATCATCTGCCTT 480  
Db 735 GTTGTGTGGGGCTGTCAAGTTCATCTCCAGCTCAACTTTTGTCTTCAACCAAAATAC 794  
QY 481 GTTGTGTGGGGCTGTCAAGTTCATCTCCAGCTCAACTTTTGTCTTCAACCAAAATAC 540  
Db 795 AACACCCAGGCGAGGTGTGTGAACCAAGTACCAAGTGTCTGGAGCCCATCAGG 854  
QY 541 AACACCCAGGCGAGGTGTGTGAACCAAGTACCAAGTGTCTGGAGCCCATCAGG 600  
Db 855 TGGAGCTGCTGATGTGGGGCTTGGAGTCTCTTGGTCTCTTTATCCCTTTGATGTTT 914  
QY 601 TGGAGCTGCTGATGTGGGGCTTGGAGTCTCTTGGTCTCTTTATCCCTTTGATGTTT 660  
Db 915 ATGATATTTTGTACAGCTTCATTTGCAAACTTGGTCAAGCTCAGAAATCTAAAGG 974  
QY 661 ATGATATTTTGTACAGCTTCATTTGCAAACTTGGTCAAGCTCAGAAATCTAAAGG 720  
Db 975 CACAAAGCCATCCGTGTAAATCATAGCTGTGGTGTGTCTTGTCTGTGTGATTCCT 1034  
QY 721 CACAAAGCCATCCGTGTAAATCATAGCTGTGGTGTGTCTTGTCTGTGTGATTCCT 780  
Db 1035 CATACATGGTCTCTGTGTGAGGCTGCAAAATTTGGTAAATGAACCGATCCTGCCAG 1094  
QY 781 CATACATGGTCTCTGTGTGAGGCTGCAAAATTTGGTAAATGAACCGATCCTGCCAG 840  
Db 1095 AGCGAAAGCTAATGGCTATACGAAAGCTGTACAGAAAGTCTGCTGCTTCCCTGCACTGC 1154  
QY 841 AGCGAAAGCTAATGGCTATACGAAAGCTGTACAGAAAGTCTGCTGCTTCCCTGCACTGC 900  
Db 1155 TGCCTGAACCTGTCTCTACGCTTTTATTTGGGAGAGTTCAGAAACTTCTTCTGAAG 1214  
QY 901 TGCCTGAACCTGTCTCTACGCTTTTATTTGGGAGAGTTCAGAAACTTCTTCTGAAG 960  
Db 1215 ATCTTGAAGGACCTGTGTGTGTGAGAGGAGTACAGTCTCAGGCTTCTCTGTGCC 1274  
QY 961 ATCTTGAAGGACCTGTGTGTGTGAGAGGAGTACAGTCTCAGGCTTCTCTGTGCC 1020  
Db 1275 GGGAGTACTCAGAAACATTTCTCGGAGACGAGTACAGACGCGAGATACGACATGCG 1334  
QY 1021 GGGAGTACTCAGAAACATTTCTCGGAGACGAGTACAGACGCGAGATACGACATGCG 1080  
Db 1335 TCGTCTTCACTATGTATAGAAAGCTGAGTCTCCCTAA 1373  
QY 1081 TCGTCTTCACTATGTATAGAAAGCTGAGTCTCCCTAA 1119

RESULT 5  
ID US-09-023-655-970 STANDARD; DNA; UNC; 1137 BP.  
AC xxxxxx  
DT  
DE Sequence 970, Application US/09023655  
CC Sequence 970, Application US/09023655  
CC GENERAL INFORMATION:  
CC APPLICANT: Cocks, Benjamin G.  
CC APPLICANT: Susan G. Stuart  
CC APPLICANT: Jeffrey J. Selhammer  
CC TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
CC TITLE OF INVENTION: EXPRESSION  
CC NUMBER OF SEQUENCES: 1508  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
CC STREET: 3174 PORTER DRIVE  
CC CITY: PALO ALTO  
CC STATE: CALIFORNIA  
CC COUNTRY: USA  
CC ZIP: 94304  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CC CURRENT APPLICATION DATA: US/09/023,655  
CC FILING DATE: HERewith  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Zeller, Karen J.  
CC REGISTRATION NUMBER: 37,071  
CC REFERENCE/DOCKET NUMBER: PA-0001 US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (650) 855-0555  
CC TELEFAX: (650) 845-4166  
CC INFORMATION FOR SEQ ID NO: 970:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1137 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC IMMEDIATE SOURCE:  
CC LIBRARY: GENBANK  
CC CLONE: g1515434  
CC SEQ SEQUENCE 1137 BP; 263 A; 270 C; 266 G; 338 T; 0 OTHER.

Query Match 95.6%; Score 1070; DB 23; Length 1137;  
Best Local Similarity 98.6%; Pred. No. 0.00e+00;  
Matches 1095; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

Db 29 ATGTTTTCGACTCCAGTGAAGATTATTTGTGTCAGTCAATFACTTCAATFACTCAGTTG 88  
QY 1 ATGTTTTCGACTCCAGTGAAGATTATTTGTGTCAGTCAATFACTTCAATFACTCAGTTG 60  
Db 89 ATTCTGAGATGTACTGTCTCTCTGTCAGGAGTTCAGGAGTTCCTCCAGGCTATTTGTAC 148  
QY 61 ATTCTGAGATGTACTGTCTCTCTGTCAGGAGTTCAGGAGTTCCTCCAGGCTATTTGTAC 120  
Db 149 CGA-TTCCCTTACTCTGTCTCTTCCCTCTGCGGGAATATTCTGCTGCTGATC 207  
QY 121 CGAATGCCCTACTCTGTCTGTCTGTCCTCTGCGGGAATATTCTGCTGCTGATC 180  
Db 208 ACCTTTGCTTTTATATAAGAGCCAGGCTCTATGACAGACGCTCTATCTCGTGAACATGGCC 267  
QY 181 ACCTTTGCTTTTATATAAGAGCCAGGCTCTATGACAGACGCTCTATCTCTTGAACATGGCC 240  
Db 268 ATTGCAGACATCTCTTTGTTTACTCTCCCATCTGGGCGAGTGAATGCCACTGGT 327  
QY 241 ATTGCAGACATCTCTTTGTTTACTCTCCCATCTGGGCGAGTGAATGCCACTGGT 300  
Db 328 GCGTGGGTTTTTCAGCAATGCCAGTGCAGTGTCTGCTAAAGGCATCTATGCCATCACTTT 387  
QY 301 GCGTGGGTTTTTCAGCAATGCCAGTGCAGTGTCTGCTAAAGGCATCTATGCCATCACTTT 360  
Db 388 AACTGCGGGATGCTGCTCTGACTTGTGATGATGACCGGTACATCGCCATTTGTACAG 447  
QY 361 AACTGCGGGATGCTGCTCTGACTTGTGATGATGACCGGTACATCGCCATTTGTACAG 420  
Db 448 GCGACTAAGTCAATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTT 507  
QY 421 GCGACTAAGTCAATTCGGGCTCCGATCCAGAACACTACCGCGCAGCAAAATCATCTGCTT 480  
Db 508 GTTGTGTGGGGCTGTGTCAGTCAATCTCCAGCTCTCTTTTGTCTTCAACCAAAATAC 567  
QY 481 GTTGTGTGGGGCTGTGTCAGTCAATCTCCAGCTCTCTTTTGTCTTCAACCAAAATAC 540  
Db 568 AACACCCCTTGGCAGGATGTGTGAACCAAGTACAGACTGTCTCGGTCGCCATCAGG 627  
QY 541 AACACCCAGGCGAGTGTGTGAACCAAGTACAGACTGTCTCGGTCGCCATCAGG 600  
Db 628 TGGAAAGCTGTGATGTTGGGGCTTGAGCTACTCTTTGTTCTTTTCCCTTTGATGTTT 687







```

CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/569,578
CC FILING DATE: 08-DEC-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: .1182.001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-601-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 578 BP; 37 A; 11 C; 62 G; 65 T; 403 OTHER.
CC
CC Query Match 8.6%; Score 96; DB 18; Length 578;
CC Best Local Similarity 12.7%; Pred. No. 2,29e-80;
CC Matches 72; Conservative 264; Mismatches 226; Indels 7; Gaps
CC
D6 3 SARVTSRSDGRVRGANNMYSKSTSSDDKVKAHGSHRSVTVKASGKKMKVDVSGKKSKD 62
C6 792 GACCATGTATTAGGAATCTGCACAGCAGAAACACACAGCATATGA-TTACAC 734
D6 63 GSDRYSHKKSKKNVTKRKYVADSKRKGCMKNKHS-DMTGTWNMGNAKTSWSKGKT 121
C6 733 GGATGGCTTTGTGCTTTTGAATCTGAGCTTGCACCAAGGTTTGAACATGAACGTGT 674
D6 122 RDDSADYHDYVGTDSWKHKSVTVKTVAHTNWRVAKHNRSCHTDNVGTGANTGNKGA 181
C6 673 AACAAATATCATGAACATCAAGGGAATAAGAACACCAAGAGTAGCTCAAGCCCCA 614
D6 182 VSMNGTSVNSHTSGSK-KRRNNYNMRAGDKKSNTRHFWGDNVRYVDWTWTKKYADSHDT 240
C6 613 TCAGCAGCTTCCACTGATGGGCTCCGAGACAGATTGGTACTTGGGTTCCACAGACATCG 554
D6 241 RRKYVHTATYRTRDKYATKATGMKYNWCDRVWKSIVHVVCVSGSTSDMTSDHSVATA 300
C6 553 TGCCTTGGGTGTGTATTTTGGTTGAAGACAAAGTAGTGCTGGGTGAGATGATGACTGAC 494
D6 301 GVTSVKNGGTVDSGRGVATYTKSTKTHSSCSVKSGNSGVYVKGPKKSDYDHS-KSSDSD 359
C6 493 GCCCCACACAAACGACGATGATTTGCTGCGCGGTAGTGTCTGATCGGAGCCGGA 434
D6 360 SYGCGARATYVTHHGTGHGKTSKTRKYDRDSSGKTKSTSHDMKWVTSRACSGST 419
C6 433 ATGACTTAGTCGCTGACAAATGGCGATGACCGGTCCCATGCTAATGCAAGTCAGGAGCA 374
D6 420 NNYMGVGMHVKTSDTAWSYDKDSGCRGSTGSKSTANRGRTSRSDGKNAGDPTKNY 479
C6 373 GCATCCCGCAGTTAAAGT-TGATGGCATAGATGCTTTTACCACTTCCACGCTGGCAT 315
D6 480 SSSKARXDKMRKGSVTVKAADRGKGKVARRSCTSSSAGRAAGDKSKKKTVSSAVAK 539
C6 314 CTGAAACCCACGACAGTGGCATGA-CTCATGTCGCCAAGATGGGAGGTAAGAACAA 256
D6 540 KRSNTTTRVKSAYVHSKGRDYNTHHG 568
C6 255 GAGCATGT-CTGCAATGGCCATGTTCAAG 228

```







[illegible]

CC APPLICANT: Lodish, Harvey F.  
CC TITLE OF INVENTION: Tgf- Type Receptor cDNAs Encoded  
CC CC  
CC TITLE OF INVENTION: Products and Uses Therefor  
CC CC  
CC NUMBER OF SEQUENCES: 8  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Hamilton Brook, Smith & Reynolds, P.C.  
CC

APPLICANT: Lodish, Harvey F.  
 TITLE OF INVENTION: TGF-Beta Type Receptor cDNAs Encoded  
 TITLE OF INVENTION: Products and Uses Therefor  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Milltia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: U.S.A.  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/311,703  
 FILING DATE:  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/786,063  
 FILING DATE: 31-OCT-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Granahan, Patricia  
 REGISTRATION NUMBER: 32,227  
 REFERENCE/DOCKET NUMBER: WH191-09  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-861-6240  
 TELEFAX: 617-861-9540  
 INFORMATION FOR SEQ ID NO: 8:







CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/446,937  
CC FILING DATE:  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/311,703  
CC FILING DATE: 23-SEP-1994  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 07/786,063  
CC FILING DATE: 31-OCT-1991  
CC CLASSIFICATION: 530  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Granahan, Patricia  
CC REGISTRATION NUMBER: 32,227  
CC REFERENCE/DOCKET NUMBER: WH191-09FW  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 617-861-6240  
CC TELEFAX: 617-861-9540  
CC INFORMATION FOR SEQ ID NO: 8:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 383 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: double  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC SEQUENCE 383 BP; 27 A; 25 C; 26 G; 34 T; 271 OTHER.

Query Match 6.8%; Score 76; DB 16; Length 383;  
Best Local Similarity 11.1%; Pred. No. 1.19e-56;  
Matches 40; Conservative 172; Mismatches 143; Indels 4; Gaps 4;

Db 16 HWTRASTHVSVNNDVMDNNGCAVCKDVRCTDKNKSMSNCSTCKVCVAVVRKNDN 75  
Cp 487 ACACAACAGGAGGAGATGTTCTCGCGGTAGTGTCTGGATCGGAGCGG-GAATGAC 429  
Db 76 TTVCHDKYHDDAASKMKKMGKFMCSNDNSNTSDVTVGSGVASVYCYVRNRKS 135  
Cp 428 TTAGTCGCTGTACATGGGATGTACCGGTCCATGTAATGCAAGTCAGGAGCGATC 369  
Db 136 STWTGTRKMSHCADRSSTSCANNHNHTDVGKRAVYKAKNTSTVAVKYDYHYSWK 195  
Cp 368 CGCAGTTAAAGTTGATGCATAGATGCCCTTTTAGCAACTT-GCACGTGGCATGCTGAA 310  
Db 196 DRKDSNKNTARKTKYKTAHAKNYTRHVSNDNRNVGVSARSQSHSDHTCGRKMVRDK 255  
Cp 309 AACCACGACGACGATGGCATGCTACTGCCAGATGGGAGAGTAAGAACAAAGAGAT 250  
Db 256 SSVNNDTCCCGSRGYSVDDANSVGTARYMAY-SRMNASKTDVYSMAVWMTSRCNA 314  
Cp 249 GPTCGAATGGCCATGTTCAAGAGATAGACGTCTGTCTATAGACCTGGCCCTTCTTATAAA 190  
Db 315 VGVKDYGSVDRVSVMKNDVDRGRNSWNHGMVCTTCWDHDARTACVARSHDRSGRS 373  
Cp 189 ACARAAGGTGATCACC-ACCAGATATTCCCGAGGAGCCCAAGAACACAGATCAAGAG 132

RESULT 15  
ID US-09-023-655-1465 STANDARD; DNA; UNC; 1500 BP.  
AC xxxxxx  
DT  
DE Sequence 1465, Application US/09023655  
CC Sequence 1465, Application US/09023655  
CC GENERAL INFORMATION:  
CC APPLICANT: Cocks, Benjamin G.  
CC APPLICANT: Susan G. Stuart  
CC APPLICANT: Jeffrey J. Seilhamer  
CC TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
CC TITLE OF INVENTION: EXPRESSION  
CC NUMBER OF SEQUENCES: 1508  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
CC STREET: 3174 PORTER DRIVE

CC CITY: PALO ALTO  
CC STATE: CALIFORNIA  
CC COUNTRY: USA  
CC ZIP: 94304  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/023,655  
CC FILING DATE: HEREWITH  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Zeller, Karen J.  
CC REGISTRATION NUMBER: 37,071  
CC REFERENCE/DOCKET NUMBER: PA-0001 US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (650) 855-0555  
CC TELEFAX: (650) 845-4166  
CC INFORMATION FOR SEQ ID NO: 1465:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1500 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC IMMEDIATE SOURCE:  
CC LIBRARY: GENBANK  
CC CLONE: 9673391  
CC SEQUENCE 1500 BP; 324 A; 457 C; 363 G; 356 T; 0 OTHER.

Query Match 6.7%; Score 75; DB 23; Length 1500;  
Best Local Similarity 63.7%; Pred. No. 1.74e-55;  
Matches 221; Conservative 0; Mismatches 122; Indels 4; Gaps 3;

Db 483 TGTGCTCAAGAAGGACGTCGGGAACCTT-TAAAGCTGTGTTCTCCCTCATCATGTACTCC 541  
Qy 76 TGTGCTCTTCGAGGAGTCAGGAGTTCCTCAGGCTATTTGTACCGAATTCCTACTCC 135  
Db 542 ATCAATTTGTTGTTGGGCTACTGGGCAATGGCTGGTGGTGTGACCTATATCTATTC 601  
Qy 136 TTGATCTGTGTTGTCCTCTGGGGAATATCTGTTGTTGATCACCTTTCCTTTTAT 195  
Db 602 AAGAGGCTCAAGACCATCAGCATACCTGCTCAACCTGGCGGTGGCAGACATCCTC 661  
Qy 196 AAGAAGGCCAGGTCTATGACAGACGTCTATCTTTGAACATGGCCATTCGACATCCTC 255  
Db 662 TTCTCTCCGACCTTCCTCTCTGGGCTTACAGCGGCGCAA--GTCC-TGGGCTTCGCT 718  
Qy 256 TTTGTTCTTACTCTCCCAATTCGAGCATGAGTATGACCTGGTGGTGGTTCAGC 315  
Db 719 GTCCACTTTTGAAGCTCATCTTTGCCATCTACAGATGAGCTTCTTTCAGTGGCATGCTC 778  
Qy 316 AATGCCAGTGCAGTTGCTTAAAGGCATCTATGCCATCAACTTTAACTGCGGAGTCTG 375  
Db 779 CTACTTCTTTGCATCAGCATGACCGCTACGTGGCCATCGTCCAGGC 825  
Qy 376 CTCCTGACTTGCATTAGCATGACCGGTACATCGCCATTGTACAGGC 422

Search completed: Sat Nov 14 01:13:28 1998  
Job time : 2043 secs.



\*\*\*\*\*

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd.

protein - protein database search, using Smith-Waterman algorithm  
Mpsrch\_pp

```
Run on:      Fri Nov 13 12:01:32 1998;  MasPar time 11.52 Seconds
           512.648 Million cell updates/sec
Tabular output not generated.
```

```
>US-08-887-977-10
Title:
Description:
Percept Score: (1-365) from US0887977.pep
                2779
Sequence: 1 MFSTPVKILCOILHITQL.....NITSROTSETADNDNASSFTM 365
```

Scoring table: PAM 150  
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

Database:

a-genseq32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29

Statistics: Mean 34.572; Variance 151.305; scale 0.228

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		ID	Description	Pred. No.
		Match	Length			
1	2779	100.0	365	29	Human dendritic cell	1.21e-355
2	990	35.6	355	15	Recombinant high affi	3.34e-81
3	990	35.6	355	6	Sequence in a high af	3.34e-81
4	984	35.4	350	12	Interleukin-8 recepto	1.26e-80
5	984	35.4	350	15	Recombinant high affi	1.26e-80
6	984	35.4	350	15	Interleukin 8 recepto	1.26e-80
7	983	35.4	350	5	Interleukin-8 recepto	1.57e-80
8	984	35.4	358	15	IL8-R type 1-GBP 130	1.26e-80
9	965	34.7	358	15	Recombinant high affi	8.48e-79
10	961	34.6	358	11	Partial sequence of s	2.05e-78
11	961	34.6	410	11	Putative seven transm	2.05e-78
12	958	34.5	378	11	Putative seven transm	3.99e-78
13	951	34.2	355	7	Human IL-8 receptor f	1.88e-77
14	951	34.2	360	6	Sequence in a low affi	1.88e-77
15	951	34.2	360	15	Interleukin 8 recepto	1.88e-77
16	951	34.2	1064	14	IL8-R type 2-GBP 130	1.88e-77
17	948	34.1	359	11	Seven transmembrane r	3.65e-77
18	927	33.4	360	15	Recombinant high affi	3.79e-75



Db 61 tfafykkarsmtvdyllnmaiaidllvltlpfawshatgawvfnscatckllkciyaiinf 120  
 QY 61 TFAYFKKARSMTDYLLNMAIADLLVLTLPFAVSHATGAWVFNATCKLLKCIYAINF 120  
 Db 121 ncgmlltciismdryiaivqatksfrlrsrtlprskliiclvvlgslvliisstvfqnky 180  
 QY 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPRSKLIICLVVGLSVLIISSTVFQNKY 180  
 Db 181 ntqgsdvcepkqyvtsepirwklmgllellffipmfmficytftvktlvqaqskr 240  
 QY 181 NTQGSDVCEPKQYVTSPIRWKLLMGLLELLFFIPMFIMFICYTFTVKTlvQAQSKR 240  
 Db 241 hkairvliavvlflacqiphnmvllvtaanlqkmrscsekligytkvtevlafhc 300  
 QY 241 HKAIRVLIIVLFLACQIPHNMVLLVTAANLQKMRSCSEKLGITYKTVTEVLAFHC 300  
 Db 301 clnplyafiqgkfrnyflkildwvrrkysgscagrysenisrqtsetadndna 360  
 QY 301 CLNPVLYAFIQGKFRNYFLKILDWCVRRKYSGSCAGRYSENISRQTSETADNDNA 360  
 Db 361 ssftm 365  
 QY 361 SSFTM 365

## RESULT 2

ID R80950 standard; Protein; 355 AA.  
 AC R80950;  
 DT 24-APR-1996 (first entry)  
 DE Recombinant high affinity interleukin-8 receptor subtype A.  
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;  
 KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;  
 KW systemic necrotising vasculitis; psoriasis; asthma; allergy; ARDS;  
 KW adult respiratory distress syndrome; neutrophil detection.  
 OS Oryctolagus cuniculus.  
 PN W09525126-A1.  
 PD 21-SEP-1995.  
 PF 09-MAR-1995; U03032.  
 PR 15-MAR-1994; US-210250.  
 PR 02-MAY-1994; US-237937.  
 PA (REPK ) REPLIGEN CORP.  
 PA (UYBO-) UNIV BOSTON.  
 PI Greenfield EA, Iarosa GJ, Navarro J, Thomas KM;  
 PI Witt DP;  
 DR WPI; 95-336945/43.  
 DR N-PSDB; Q99949.  
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -  
 PT useful for treating inflammatory disorders, for detecting  
 PT neutrophil(s) and for isolating IL-8 receptor from liq.mixt.  
 PS Claim 2; Fig 1A-B; 74pp; English.  
 CC Monoclonal antibodies were raised against recombinant interleukin-8  
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources  
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8A) is  
 CC a high affinity receptor and the B subtype receptor (IL-8B) is a  
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to  
 CC the IL-8 binding domain thus blocking its activation. The mAbs are  
 CC useful for treating inflammatory disorders (see key words) and for  
 CC detecting the presence of neutrophils in a biological sample. The  
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.  
 SQ Sequence 355 AA;

Query Match 35.6%; Score 990; DB 15; Length 355;  
 Best Local Similarity 45.1%; Pred. No. 3.34e-81;  
 Matches 130; Conservative 78; Mismatches 70; Indels 10; Gaps 7;

Db 45 yvvviyalvflilgslvmlvlysnrsrvtdvyllynlamadllfaltmpiwavsk 104  
 QY 45 YLVRIAYSLICVLGLGNLWVITFAFYKARSMTDYLNNMAIADLLVLTLPFWAVSH 97  
 Db 105 ekg-wifgtplckvslvkevnfygilllacisvdrlyaihvhatrltqk-rhl--vkf 160  
 QY 98 ATGAWVFNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATKSFRLSRRTLPRSKI 157

Db 161 iclgiwalisllslpfflgrqvfspnnssprvc---yedghntakwrmvrlilphtfgfi 217  
 QY 158 ICLVWGLSVLIISSTVFVNQKYNQTQSD-VCEPKYQIVS-EPIRWKLLMGLLELFGFF 215  
 Db 218 lpllvmlfcygtlrtlfaqhmgqkhramrvfawvllfllcwlpynlvlladtlmrthv 277  
 QY 216 IPLMFIMFCYTFIVKTLVQAQNSKRHRKAIRVIAVVLFLACQIPHNNVLLVTAANLGM 275  
 Db 278 iqetccqrndidraldateilgflhscinpliyafiqgnfngfklml 325  
 QY 276 NR-SCQSEKLGITYKTVTEVLAFHLHCLLPVLYAFIQGKFRNYFLKIL 322

## RESULT 3

ID R28272 standard; Protein; 355 AA.  
 AC R28272;  
 DT 04-APR-1993 (first entry)  
 DE Sequence in a high affinity recombinant rabbit interleukin-8  
 DE (IL-8) receptor polypeptide in F3R.  
 KW IL-8 receptor polypeptide; G-protein-coupled receptor.  
 OS Oryctolagus cuniculus.  
 PN W09218641-A.  
 PD 29-OCT-1992.  
 PF 10-APR-1992; U02977.  
 PR 10-APR-1991; US-685101.  
 PR 09-JUL-1991; US-726606.  
 PR 09-DEC-1991; US-803842.  
 PA (REPK ) REPLIGEN CORP.  
 PA (UYBO-) UNIV BOSTON.  
 PI Navarro J, Thomas KM, Witt DP;  
 PI WPI; 92-382123/46.  
 DR N-PSDB; Q30011.  
 PT Recombinant mammalian interleukin-8 receptor - used for screening  
 PT interleukin-8 binding antagonists, used to treat inflammation  
 PS Claim 2; Fig 1; 71pp; English.  
 CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit  
 CC peritoneal neutrophils and used as a source of poly(A)+ RNA, to  
 CC produce a rabbit neutrophil cDNA library. 250,000 recombinant  
 CC plaques were screened for those which hybridized to an antisense  
 CC oligonucleotide (Q30015). This probe was designed based on the  
 CC sequence derived from the second transmembrane domain of G-protein-  
 CC coupled receptors. After tertiary screening, six plaques were  
 CC isolated. The insert of one of these plaques, termed F3R was of 2.5  
 CC kb in size. This insert was sequenced. The protein deduced from  
 CC the F3R clone demonstrates that it belongs to the family of  
 CC G-protein-coupled receptors. The deduced protein sequence  
 CC indicates seven putative transmembrane segments.  
 SQ Sequence 355 AA;

Query Match 35.6%; Score 990; DB 6; Length 355;  
 Best Local Similarity 45.1%; Pred. No. 3.34e-81;  
 Matches 130; Conservative 78; Mismatches 70; Indels 10; Gaps 7;

Db 45 yvvviyalvflilgslvmlvlysnrsrvtdvyllynlamadllfaltmpiwavsk 104  
 QY 45 YLVRIAYSLICVLGLGNLWVITFAFYKARSMTDYLNNMAIADLLVLTLPFWAVSH 97  
 Db 105 ekg-wifgtplckvslvkevnfygilllacisvdrlyaihvhatrltqk-rhl--vkf 160  
 QY 98 ATGAWVFNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATKSFRLSRRTLPRSKI 157  
 Db 161 iclgiwalisllslpfflgrqvfspnnssprvc---yedghntakwrmvrlilphtfgfi 217  
 QY 158 ICLVWGLSVLIISSTVFVNQKYNQTQSD-VCEPKYQIVS-EPIRWKLLMGLLELFGFF 215  
 Db 218 lpllvmlfcygtlrtlfaqhmgqkhramrvfawvllfllcwlpynlvlladtlmrthv 277  
 QY 216 IPLMFIMFCYTFIVKTLVQAQNSKRHRKAIRVIAVVLFLACQIPHNNVLLVTAANLGM 275  
 Db 278 iqetccqrndidraldateilgflhscinpliyafiqgnfngfklml 325  
 QY 276 NR-SCQSEKLGITYKTVTEVLAFHLHCLLPVLYAFIQGKFRNYFLKIL 322



RESULT 4  
 ID R68811 standard; Protein; 350 AA.  
 AC R68811;  
 DT 18-JUL-1995 (first entry)  
 DE Interleukin-8 receptor.  
 KW Interleukin-8 receptor.  
 KW platelet factor superfamily receptor; neutrophil; chemotactic;  
 KW inflammation; inflammatory disease; arthritis; emphysema; cystic;  
 KW fibrosis; colitis; bronchitis; meningitis; therapeutic.  
 OS Homo sapiens.  
 PN W09428931-A.  
 PD 22-DEC-1994.  
 PF 07-JUN-1994; U06380.  
 PR 11-JUN-1993; US-076093.  
 FA (GETH) GENENTECH INC.  
 PI Chuntharapai A, Hebert C, Kim KJ, Lee J;  
 DR WPI: 95-036114/05.  
 DR N-PSDB: Q80520.  
 PT Treatment of inflammatory disorders - by administering an  
 PT antibody capable of binding a platelet factor 4 superfamily  
 PT receptor polypeptide  
 PS Disclosure; Page 51-54; 83pp; English.  
 CC A cDNA library constructed from human neutrophil mRNA in pRK5B was  
 CC transfected into COS-7 cells, and the cells were screened with 125I-  
 CC IL-8. The DNA sequence of isolated cDNA clone pRK5B.118r1.1  
 CC encoding human IL-8 receptor, is given in Q80520 and the predicted  
 CC amino acid sequence in R68811. The receptor is used to raise  
 CC antibodies that neutralize the activity of PF4AR, e.g. IL-8 receptor.  
 SQ Sequence 350 AA;  
  
 Query Match 35.4%; Score 984; DB 12; Length 350;  
 Best Local Similarity 44.4%; Pred. No. 1.26e-80;  
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;  
  
 Db 27 y-spcmletetinkvviayalvllsllgnsllmvllysvgrsvtdvylalnlaad 85  
 QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGILGNLVITFAFYKARSMTDVLNMAIAD 83  
  
 Db 86 llfaltlpwaaskvng-wifgtfclckvsvllkevnyfsgilllacisvdrlyalvhatr 144  
 QY 84 ILFVLTLPWAVSHATGAWFVSNTCKLLKGIYAINFNCGMLLTCISMDRYIAIVQATK 143  
  
 Db 145 tltoqk-rhl--vkfvclogswlmslslpflfrqayhpnssvpc---yevlgndtakw 198  
 QY 144 SPRLRSRLPRSKIIICLVVWGLSVIISSTFVFNQKYNTOGSD-VCEPKYQTVS-EP 201  
  
 Db 199 rmvlrplhtgfivplfvmficygftlrlfkahmgqkhrmvifavvllfllcwlp 258  
 QY 202 KLLMLGLELLFGFFPLMFIMFCYTFIVKTLVQAQNSRHRKAIIVIAVFLACQIPH 261  
  
 Db 259 nlvlladtlmrtqvigetccerrnngiraldatelilgflhscnlpiyafiqnfrhgfik 318  
 QY 262 NNVLLVTAANLGKMR-SCQSEKLIQYTKVTVEVLAFLHCLNPNVLYAFIQGFRNFK 320  
  
 Db 319 il 320  
 QY 321 IL 322  
  
 RESULT 5  
 ID R80951 standard; Protein; 350 AA.  
 AC R80951;  
 DT 24-APR-1996 (first entry)  
 DE Recombinant high affinity interleukin-8 receptor subtype A.  
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;  
 KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;  
 KW systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;  
 KW adult respiratory distress syndrome; neutrophil detection.  
 OS Homo sapiens.  
 PN W09525126-A1.  
 PD 21-SEP-1995.  
 PF 09-MAR-1995; U03032.

PR 15-MAR-1994; US-210250.  
 PR 02-MAY-1994; US-237937.  
 PA (REPK) REPLIGEN CORP.  
 PA (UYBO-) UNIV BOSTON.  
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;  
 PI Witt DT;  
 DR WPI: 95-336945/43.  
 DR N-PSDB: Q99950.  
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -  
 PT useful for treating inflammatory disorders, for detecting  
 PT neutrophil(s) and for isolating IL-8 receptor from liq.mixt.  
 PS Disclosure; Fig 2A-B; 74pp; English.  
 CC Monoclonal antibodies were raised against recombinant interleukin-8  
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources  
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8A) is  
 CC a high affinity receptor and the B subtype receptor (IL-8B) is a  
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to  
 CC the IL-8 binding domain thus blocking its activation. The mAbs are  
 CC useful for treating inflammatory disorders (see key words) and for  
 CC detecting the presence of neutrophils in a biological sample. The  
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.  
 SQ Sequence 350 AA;  
  
 Query Match 35.4%; Score 984; DB 15; Length 350;  
 Best Local Similarity 44.4%; Pred. No. 1.26e-80;  
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;  
  
 Db 27 y-spcmletetinkvviayalvllsllgnsllmvllysvgrsvtdvylalnlaad 85  
 QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGILGNLVITFAFYKARSMTDVLNMAIAD 83  
  
 Db 86 llfaltlpwaaskvng-wifgtfclckvsvllkevnyfsgilllacisvdrlyalvhatr 144  
 QY 84 ILFVLTLPWAVSHATGAWFVSNTCKLLKGIYAINFNCGMLLTCISMDRYIAIVQATK 143  
  
 Db 145 tltoqk-rhl--vkfvclogswlmslslpflfrqayhpnssvpc---yevlgndtakw 198  
 QY 144 SFRUSRLTPRSKIIICLVVWGLSVIISSTFVFNQKYNTOGSD-VCEPKYQTVS-EP 201  
  
 Db 199 rmvlrplhtgfivplfvmficygftlrlfkahmgqkhrmvifavvllfllcwlp 258  
 QY 202 KLLMLGLELLFGFFPLMFIMFCYTFIVKTLVQAQNSRHRKAIIVIAVFLACQIPH 261  
  
 Db 259 nlvlladtlmrtqvigetccerrnngiraldatelilgflhscnlpiyafiqnfrhgfik 318  
 QY 262 NNVLLVTAANLGKMR-SCQSEKLIQYTKVTVEVLAFLHCLNPNVLYAFIQGFRNFK 320  
  
 Db 319 il 320  
 QY 321 IL 322  
  
 RESULT 6  
 ID R80756 standard; Protein; 350 AA.  
 AC R80756;  
 DT 26-MAR-1996 (first entry)  
 DE Interleukin 8 receptor A partial sequence.  
 KW Interleukin; IL-8; inflammation; psoriasis; dermatitis;  
 KW rheumatoid arthritis; inflammatory bowel disease;  
 KW chronic lung inflammation; treatment; antibody;  
 KW affinity purification; detection.  
 OS Homo sapiens.  
 PN US5440021-A.  
 PD 08-AUG-1995.  
 PF 29-MAR-1991; 677211.  
 PR 29-MAR-1991; US-677211.  
 PR 25-FEB-1994; US-202056.  
 PA (CHUN)/ CHUNTHARAPAI A.  
 PA (HEBE)/ HEBERT C.  
 PA (KIMK)/ KIM K J.  
 PA (LEEJ)/ LEE J.  
 PI Chuntharapai A, Hebert C, Kim KJ, Lee J;  
 DR WPI: 95-283151/37.



DR N-PSDB; Q99006.  
PT New antibodies against interleukin 8 type B receptor - used to treat  
PT or prevent inflammation, also for detecting receptor expression and  
PT purification.  
PS Example 2; Columns 41-44; 62pp; English.  
CC Antibodies directed against the interleukin-8 receptor B can be used  
CC to treat or prevent inflammation e.g. psoriasis, dermatitis,  
CC rheumatoid arthritis and particularly inflammatory bowel disease and  
CC chronic lung inflammation. When immobilised, these antibodies may  
CC be used to detect interleukin-8 receptor B expression in cells and  
CC tissues and for affinity purification of interleukin-8 receptor B  
CC from cells. A sequence encoding amino acids 23-314 of this fragment  
CC of the high affinity interleukin-8 receptor, was used to probe lambda  
CC cgt10 cDNA libraries for the human interleukin-8 type B receptor.  
SQ Sequence 350 AA;

Query Match 35.4%; Score 984; DB 15; Length 350;  
Best Local Similarity 44.4%; Pred. No. 1.26e-80;  
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 y-spcmetetlnkyvviayalvfillgnslnvmlvilyrsvrtdvylalnala 85  
QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGLNGLVITFAFYKARSMVDVLLNMAIAD 83

Db 86 llfaltlpiwaaskvng-wifgtfclkvsvllkevnfygilllacisvdrlylaivhatr 144  
QY 84 ILFVLTLPFWAVSHATGAWVFSNATCKLKGIAINFNCGMILLTCISMDRYIAIVQATK 143

Db 145 tltdq-rhl--vkfvcgclgwmnslpfflfrqayhpnsspyc---yevlgndtakw 198  
QY 144 SFRLRSRTLPKSKIIICLVWGLSVIISSTFWFNQKYNTOGSD-VCEPKYQTVS-EP1RW 201

Db 199 rmvlriltphgtfivplvmlfcygtftrtlfkahmgqkhrmvifavvllfclwlp 258  
QY 202 KLLMLGLELLFGFFIPLMFMEFCTYFIVKTLVQAQNSKRHKAIIRVIAVVLVFLACQIPH 261

Db 259 nlvlladtlnrtqvietcennnigraldateilgfhscnlpiyafignfrhfglk 318  
QY 262 NMVLLVTAANLGRMNR-SCOSEKLIGYTKTVTEVLAFLHCLLPVLYAFIGQFRNYFLK 320

Db 319 il 320  
QY 321 IL 322

RESULT 7  
ID R27791 standard; Protein; 350 AA.  
AC R27791;  
DT 12-MAR-1993 (first entry)  
DE Interleukin-8 receptor.  
KW IL-8; G-protein coupled receptor family; rhodopsin superfamily;  
KW pro-inflammatory cytokine.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT domain 40..67  
FT /label= "transmembrane"  
FT /note= "putative"  
FT domain 76..98  
FT /label= "transmembrane"  
FT /note= "putative"  
FT domain 112..133  
FT /label= "transmembrane"  
FT /note= "putative"  
FT domain 155..174  
FT /label= "transmembrane"  
FT /note= "putative"  
FT domain 204..226  
FT /label= "transmembrane"  
FT /note= "putative"  
FT domain 243..264  
FT /label= "transmembrane"  
FT /note= "putative"  
FT domain 292..312

FT /label= "transmembrane"  
FT /note= "putative"  
PN W09217497-A.  
PD 15-OCT-1992.  
PR 23-MAR-1992; U02317.  
PR 29-MAR-1991; US-677211.  
PR 19-DEC-1991; US-810782.  
PA (GETH ) GENENTECH INC.  
PI Holmes WE, Lee J, Wood WI;  
WIPI; 92-366191/44.  
DR N-PSDB; Q29505.  
PT Isolated human platelet factor 4 super-family receptor  
PT polypeptide and corresp. antibodies and DNA - useful as  
PT diagnostic and screening agents, and for treating inflammation or  
PT PF4AR-mediated disorders  
PS Claim 7; Fig 2; 78pp; English.  
CC A cDNA library constructed from human neutrophil mRNA in the mammalian  
CC expression vector pRK5B was transfected into COS-7 cells as pools of  
CC 2500 clones. One positive pool from the first 58 transfections was  
CC partitioned into smaller pools until a pure clone (PRK5B.118r1.1)  
CC was obtained. The ORF encodes a protein of 350 amino acids which  
CC shares several features with the G-protein coupled receptors of the  
CC rhodopsin superfamily, including 7 hydrophobic (transmembrane)  
CC domains. The IL-8 receptor is a preferred PF4AR superfamily member  
CC of the invention. See also Q29506 and Q37107.  
SQ Sequence 350 AA;

Query Match 35.4%; Score 983; DB 5; Length 350;  
Best Local Similarity 44.4%; Pred. No. 1.57e-80;  
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 y-spcmetetlnkyvviayalvfillgnslnvmlvilyrsvrtdvylalnala 85  
QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGLNGLVITFAFYKARSMVDVLLNMAIAD 83

Db 86 llfaltlpiwaaskvng-wifgtfclkvsvllkevnfygilllacisvdrlylaivhatr 144  
QY 84 ILFVLTLPFWAVSHATGAWVFSNATCKLKGIAINFNCGMILLTCISMDRYIAIVQATK 143

Db 145 tltdq-rhl--vkfvcgclgwmnslpfflfrqayhpnsspyc---yevlgndtakw 198  
QY 144 SFRLRSRTLPKSKIIICLVWGLSVIISSTFWFNQKYNTOGSD-VCEPKYQTVS-EP1RW 201

Db 199 rmvlriltphgtfivplvmlfcygtftrtlfkahmgqkhrmvifavvllfclwlp 258  
QY 202 KLLMLGLELLFGFFIPLMFMEFCTYFIVKTLVQAQNSKRHKAIIRVIAVVLVFLACQIPH 261

Db 259 nlvlladtlnrtqvietcennnigraldateilgfhscnlpiyafignfrhfglk 318  
QY 262 NMVLLVTAANLGRMNR-SCOSEKLIGYTKTVTEVLAFLHCLLPVLYAFIGQFRNYFLK 320

Db 319 il 320  
QY 321 IL 322

RESULT 8  
ID R70123 standard; Protein; 1060 AA.  
AC R70123;  
DT 14-FEB-1996 (first entry)  
DE IL8-R type 1-GPB 130 fusion protein.  
KW Hybrid peptide; malaria parasite; Plasmodium falciparum; fusion protein;  
KW red blood cell; cytokine receptor; glycoprotein binding peptide 130;  
KW GBP 130; GBP; glycoprotein binding peptide homologue; glycoprotein A.  
OS Chimeric Homo sapiens.  
OS Chimeric Plasmodium falciparum.  
PN W09506737-A.  
PD 09-MAR-1995.  
PF 01-SEP-1994; G01900.  
PR 03-SEP-1993; GB-018350.  
PR 23-AUG-1994; GB-017021.  
PA (PREN/) PRENDERGAST K F.  
PI Prendergast KF;











OS Homo sapiens.  
 PN W09306229-A.  
 PD 01-APR-1993.  
 PF 14-SEP-1992; U07641.  
 PR 13-SEP-1991; US-759568.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
 PI Murphy PM;  
 DR WPI: 93-117549/14.  
 DR N-PSDB; Q38747.  
 PT New interleukin-8 receptor aminoacid sequence - and corresp. cDNA  
 expressed in Xenopus laevis oocytes or transfected host cells,  
 PT for screening ligands of IL-8 receptor and gene therapy  
 PS Claim 1; Fig 3; 39pp; English.  
 CC cDNA libraries from 2 and 3.5 kb fractions of poly(A)+ RNA from HL60  
 CC neutrophils sep'd. of a sucrose gradient were made in UniZAP. The  
 CC libraries were screened with F3R oligonucleotide probe (from rabbit  
 CC IL-8 receptor) and under low stringency with a p2 cDNA probe  
 CC synthesised from random primers, to isolated the clone p2, encoding  
 CC human IL-8 receptor protein. The presence or absence of the DNA  
 CC encoding IL-8R or related MIP-2 receptor may be detected using  
 CC portions of the p2 clone as probes. P2 may also be used to screen  
 CC for ligands of IL-8R and may also be used in gene therapy to treat a  
 CC patient deficient in IL-8R. The IL-8R is a gro receptor and has  
 CC intracellular calcium-mobilising and ligand-binding properties.  
 SQ Sequence 355 AA;

Query Match 34.2%; Score 951; DB 7; Length 355;  
 Best Local Similarity 43.8%; Pred. No. 1.88e-77;  
 Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;  
 Db 44 yfvviyalvllslgnslymvllyrvgrsvtdvllnlaladlalfaltlpiwaask 103  
 QY 38 YLXRIAYSLICVLGLGNILVITFAFYKARSMTDVLNNAIADILFVLTLPFWAVSH 97  
 Db 104 vng-wifgtfclckvsvllkevnfygillacisvdrlylaihvartltqk-ryl--vkf 159  
 QY 98 ATGAWFSNATCKLLKGIYAINFNGMLLATCISMDRYIAIVQATKSFRLSRTPRSKI 157  
 Db 160 iclsigwslslalpvllfrtvyssnvspac---yedmgntanwmlrlilpqsfghi 216  
 QY 158 ICLVWGLSVIISSTFVFNQK-YNTQSDVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215  
 Db 217 vpllimifcygftlrlfkahmgqkhravrfavvllfllcwlpylnvlladtlmrtqv 276  
 QY 216 IPLMFIMFCYTFIVKTLVQAQSKRHKRAIRVIAVFLACQIPHNVLVTAANLGKM 275  
 Db 277 iqetcermhldaldateilgilhscnlpiyafiqgkfrhglkil 324  
 QY 276 NR-SCQSEKLGKTYKTVTEVLAFLHCLLPVLYAFIGQKFRNYFLKIL 322

RESULT 14  
 ID R28273 standard; Protein; 360 AA.  
 AC R28273;  
 DT 04-APR-1993 (first entry)  
 DE Sequence in a low affinity recombinant human interleukin-8  
 DE (IL-8) receptor polypeptide in 4AB.  
 KW IL-8 receptor polypeptide; G-protein-coupled receptor.  
 OS Homo sapiens.  
 PN W09218641-A.  
 PD 29-OCT-1992.  
 PF 10-APR-1992;  
 PR 10-APR-1991; US-685101.  
 PR 09-JUL-1991; US-726606.  
 PR 09-DEC-1991; US-803842.  
 PA (REPK ) REPLIGEN CORP.  
 PA (DYBO-) UNIV BOSTON.  
 PI Navarro J, Thomas KM, Witt DP;  
 DR WPI: 92-382123/46.  
 DR N-PSDB; Q30012.  
 PT Recombinant mammalian interleukin-8 receptor - used for screening  
 PT interleukin-8 binding antagonists, used to treat inflammation  
 PS Disclosure; Fig 2; 71pp; English.

CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit  
 CC peritoneal neutrophils and used as a source of poly(A)+ RNA. to  
 CC produce a rabbit neutrophil cDNA library. 250,000 recombinant  
 CC plaques were screened for those which hybridized to an antisense  
 CC oligonucleotide (Q30015). This probe was designed based on the  
 CC sequence derived from the second transmembrane domain of G-protein-  
 CC coupled receptors. After tertiary screening, six plaques were  
 CC isolated. The insert of one of these plaques, termed F3R was of 2.5  
 CC kb in size. This insert was sequenced. The protein deduced from  
 CC the F3R clone demonstrates that it belongs to the family of  
 CC G-protein-coupled receptors. The deduced protein sequence  
 CC indicates seven putative transmembrane segments. A human  
 CC peripheral blood leukocyte lambda gt11 cDNA library (5' stretch) was  
 CC screened with a 652 bp EcoRI/BamHI fragment (including nucleotides  
 CC -27 to 825) of the rabbit F3R clone. After tertiary screening  
 CC several human clones which hybridized to the rabbit IL-8 probe  
 CC were isolated. The insert of one such clone, termed 4AB was  
 CC sequenced (Q30012).  
 SQ Sequence 360 AA;

Query Match 34.2%; Score 951; DB 6; Length 360;  
 Best Local Similarity 43.8%; Pred. No. 1.88e-77;  
 Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;  
 Db 49 yfvviyalvllslgnslymvllyrvgrsvtdvllnlaladlalfaltlpiwaask 108  
 QY 38 YLXRIAYSLICVLGLGNILVITFAFYKARSMTDVLNNAIADILFVLTLPFWAVSH 97  
 Db 109 vng-wifgtfclckvsvllkevnfygillacisvdrlylaihvartltqk-ryl--vkf 164  
 QY 98 ATGAWFSNATCKLLKGIYAINFNGMLLATCISMDRYIAIVQATKSFRLSRTPRSKI 157  
 Db 165 iclsigwslslalpvllfrtvyssnvspac---yedmgntanwmlrlilpqsfghi 221  
 QY 158 ICLVWGLSVIISSTFVFNQK-YNTQSDVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215  
 Db 222 vpllimifcygftlrlfkahmgqkhravrfavvllfllcwlpylnvlladtlmrtqv 281  
 QY 216 IPLMFIMFCYTFIVKTLVQAQSKRHKRAIRVIAVFLACQIPHNVLVTAANLGKM 275  
 Db 282 iqetcermhldaldateilgilhscnlpiyafiqgkfrhglkil 329  
 QY 276 NR-SCQSEKLGKTYKTVTEVLAFLHCLLPVLYAFIGQKFRNYFLKIL 322

RESULT 15  
 ID R80758 standard; Protein; 360 AA.  
 AC R80758;  
 DT 26-MAR-1996 (first entry)  
 DE Interleukin 8 receptor B.  
 KW Interleukin; IL-8; inflammation; psoriasis; dermatitis;  
 KW rheumatoid arthritis; inflammatory bowel disease;  
 KW chronic lung inflammation; treatment; antibody;  
 KW affinity purification; detection.  
 OS Homo sapiens.  
 PN US5440021-A.  
 PD 08-AUG-1995.  
 PF 29-MAR-1991; 677211.  
 PR 29-MAR-1991; US-677211.  
 PR 25-FEB-1994; US-202056.  
 PA (CHUN/) CHUNTHARAPAI A.  
 PA (HEBE/) HEBERT C.  
 PA (KIMK/) KIM K J.  
 PA (LEEJ/) LEE J.  
 PI Chuntharapai A, Hebert C, Kim KJ, Lee J;  
 DR WPI: 95-283151/37.  
 DR N-PSDB; Q99008.  
 PT New antibodies against interleukin 8 type B receptor - used to treat  
 PT or prevent inflammation, also for detecting receptor expression and  
 PT purification.  
 PS Disclosure; Columns 53-56; 62pp; English.  
 CC Antibodies directed against the interleukin-8 receptor B can be used  
 CC to treat or prevent inflammation e.g. psoriasis, dermatitis,



SQ Sequence 360 AA;

Search completed: Fri Nov 13 12:03:22 1998  
Job time : 110 secs.



\*\*\*\*\*  
WISREH  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Fri Nov 13 12:03:41 1998; MasPar time 16.03 Seconds  
Tabular output not generated. 831.784 Million cell updates/sec

Title: >US-08-887-977-10  
Description: (1-365) from US08887977.ppe  
Perfect Score: 2779  
Sequence: 1 MFSTPVKILLCOSILHTQL.....NISQTSETADNDNASSFTM 365

Scoring table: PAM 150  
Gap 11  
Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir56  
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d  
Statistics: Mean 47.004; Variance 114.799; scale 0.409

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2445	88.0	369	2	JC5068	G protein-coupled rec 0.00e+00
2	990	35.6	355	2	JQ1231	interleukin-8 recepto 3.32e-140
3	986	35.5	350	2	A39445	interleukin-8 recepto 1.59e-139
4	965	34.7	358	2	A53752	interleukin-8 recepto 5.81e-136
5	961	34.6	378	2	B55735	lymphocyte-specific G 2.77e-135
6	959	34.5	378	2	A53735	G protein-coupled rec 6.06e-135
7	951	34.2	360	2	A53611	interleukin-8 recepto 1.38e-133
8	927	33.4	360	2	A57160	chemokine (C-C) recep 1.60e-129
9	928	33.4	378	2	A45680	G protein-coupled pep 1.08e-129
10	906	32.6	356	2	S42096	interleukin-8 recepto 5.72e-126
11	891	32.1	359	2	A48921	interleukin-8 recepto 1.96e-123
12	870	31.3	360	2	JC4587	chemokine (C-C) recep 6.90e-120
13	839	30.2	355	2	A45177	chemokine (C-C) recep 1.16e-114
14	836	30.1	354	2	A23669	interleukin-8 recepto 3.71e-114
15	835	30.0	359	2	I49341	MIP-1 alpha receptor 5.47e-114
16	793	28.5	355	2	G02436	C-C chemokine receptor 6.29e-107
17	784	28.2	355	2	A57337	chemokine (C-C) recep 2.04e-105
18	776	27.9	352	2	A43113	chemokine (C-C) recep 4.47e-104
19	769	27.7	352	2	G02653	CC chemokine receptor 6.66e-103
20	762	27.4	355	2	I49339	macrophage inflammato 9.92e-102
21	755	27.2	360	2	JC2443	chemokine (C-C) recep 1.47e-100
22	751	27.0	370	2	I38450	chemokine (C-C) recep 6.88e-100
23	738	26.6	355	2	JC5067	G protein-coupled rec 1.03e-97

24	734	26.4	354	2	I58186	hypothetical G-protein 4.78e-97
25	726	26.1	374	2	S32785	G protein-coupled rec 1.04e-95
26	723	26.0	359	2	S15403	angiotensin II recept 3.28e-95
27	721	25.9	355	2	JC4304	orphan G-protein-coup 7.08e-95
28	717	25.8	352	2	A45747	leukocyte-derived sev 3.29e-94
29	716	25.8	359	2	A42656	angiotensin II recept 4.84e-94
30	713	25.7	359	2	JH0621	angiotensin receptor 1.53e-93
31	715	25.7	374	2	S42628	G-protein coupled rec 7.10e-94
32	711	25.6	359	2	JC1104	angiotensin II recept 3.30e-93
33	708	25.5	352	2	G00048	angiotensin II recept 1.04e-92
34	709	25.5	359	2	JQ1516	angiotensin II recept 7.11e-93
35	706	25.4	353	2	S28787	neuropeptide Y/peptid 2.25e-92
36	707	25.4	356	2	I49340	MIP-1 alpha receptor 1.53e-92
37	707	25.4	359	2	JC2134	angiotensin II recept 1.53e-92
38	699	25.2	359	2	JC1193	angiotensin II recept 3.29e-91
39	698	25.1	350	2	JN0621	G-protein coupled rec 4.83e-91
40	698	25.1	359	2	S44425	angiotensin II recept 4.83e-91
41	697	25.1	359	2	A48857	AT1 angiotensin II re 7.09e-91
42	690	24.8	359	2	JC1194	angiotensin II recept 1.04e-89
43	687	24.7	327	2	S56162	MDR15 protein - huma 3.27e-89
44	687	24.7	372	2	S26667	G protein-coupled rec 3.27e-89
45	683	24.6	383	2	S55594	G protein-coupled rec 1.51e-88

ALIGNMENTS

RESULT 1  
ENTRY JC5068 #type complete  
TITLE G protein-coupled receptor CKR-L3 - human  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 10-Sep-1997  
ACCESSIONS JC5068  
REFERENCE JC5067  
#authors Zaballlos, A.; Varona, R.; Gutierrez, J.; Lind, P.; Marquez, G.  
#journal Biochem. Biophys. Res. Commun. (1996) 227:846-853  
#title Molecular cloning and RNA expression of two new human chemokine receptor-like genes.  
#accession JC5068  
#molecule\_type DNA  
#residues 1-369 #label ZAB  
#cross-references EMBL:Z79784; NID:gl668737; PID:e264774; PID:gl668738  
COMMENT This protein belongs to the family of alpha chemokine receptors.  
KEYWORDS G protein-coupled receptor; transmembrane protein  
FEATURE  
42-68 #domain transmembrane #status predicted #label TM1\  
79-99 #domain transmembrane #status predicted #label TM2\  
115-136 #domain transmembrane #status predicted #label TM3\  
160-180 #domain transmembrane #status predicted #label TM4\  
212-233 #domain transmembrane #status predicted #label TM5\  
250-271 #domain transmembrane #status predicted #label TM6\  
292-315 #domain transmembrane #status predicted #label TM7  
#length 369 #molecular-weight 42002 #checksum 1020

Query Match 88.0%; Score 2445; DB 2; Length 369;  
Best Local Similarity 98.8%; Pred. NO. 0.00e+00;  
Matches 324; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
Db 42 LFVPDIATSLICVFGLLGNILWITFAFYKKARSMTDVYLLNMAIADILFVLTPFWAVSH 101  
QY 39 LY-RIATSLICVGLLGNILWITFAFYKKARSMTDVYLLNMAIADILFVLTPFWAVSH 97  
Db 102 ATGAWVESNATCKLLKGIYAINFCNGLMLTCTISMDRYIAIVQATKSPRLSRTPRSKI 161  
QY 98 ATGAWVESNATCKLLKGIYAINFCNGLMLTCTISMDRYIAIVQATKSPRLSRTPRSKI 157  
Db 162 ICLVWGLSVIISSTTFVFNQKYNQTCSDVCEPKYQVSEPIRWKLLMLGLLELFGFFIP 221  
QY 158 ICLVWGLSVIISSTTFVFNQKYNQTCSDVCEPKYQVSEPIRWKLLMLGLLELFGFFIP 217  
Db 222 LMFIMFYCTIVKTLVQAKSKRKAIRVIAVVLVFLACQIPHNVLVLTAAKGNR 281



```

Qy 218 LMFIFCYTFIVKTLVQAOQSKRRKRAIRVIAVLVFLACQIPHNMLVLTAAANLGMNR 277
Db 282 SCQSEKLLGYTKTVEVLAFLHCLLPVLYAFIGQFRNYFLKILKDLWCVRKRYKSSGF 341
Qy 278 SCQSEKLLGYTKTVEVLAFLHCLLPVLYAFIGQFRNYFLKILKDLWCVRKRYKSSGF 337
Db 342 SCAGRYSENISQTSQTSADNDNASSFTM 369
Qy 338 SCAGRYSENISQTSQTSADNDNASSFTM 365

RESULT 2
ENTRY #type complete
TITLE interleukin-8 receptor - rabbit
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
        rabbit
DATE 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change
ACCESSIONS JQ1231
REFERENCE JQ1231; A46483
#authors Beckmann, M.P.; Munger, W.E.; Kozlosky, C.; VandenBos, T.;
        Price, V.; Lyman, S.; Gerard, N.P.; Gerard, C.; Cerretti,
        D.P.
#journal Biochem. Biophys. Res. Commun. (1991) 179:784-789
#title Molecular characterization of the interleukin-8 receptor.
#cross-references MUID:91378994
#accession JQ1231
#molecule_type DNA
#residues 1-355 #label BEC
#cross-references GB:M74240; NID:g165438; PID:g165439
REFERENCE A46483
#authors Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
#journal J. Immunol. (1992) 148:1261-1264
#title Characterization of complementary DNA clones encoding the
        rabbit IL-8 receptor.
#cross-references MUID:92148149
#accession A46483
#status preliminary
#molecule_type mRNA
#residues 1-355 #label LEE
#cross-references GB:M82873; NID:g165440; PID:g165441
#experimental_source neutrophils
#note sequence extracted from NCBI backbone (NCBIN:81526,
        NCBI:P:81530)
KEYWORDS G protein-coupled receptor; transmembrane protein
SUMMARY #length 355 #molecular-weight 40622 #checksum 102

Query Match 35.6%; Score 990; DB 2; Length 355;
Best Local Similarity 45.1%; Pred. No. 3.32e-140;
Matches 130; Conservative 78; Mismatches 70; Indels 10; Gaps 7;

Db 45 YVVVYIYALVELLSILGNSLVMLVLYSRNSRVTVDVYLLNLAADLFLALPMPWAVSK 104
Qy 38 YLYRIAYSLICVLGLNLIYVITFAFKKARSMTDVLNNAIADILFLVLPFWAVSH 97
Db 105 EKG-WIFGTPLCVSVLYKVENYFGSILLACISVDYRLAIYHATRTQK-RHL--VKF 160
Qy 98 ATGAWFVSNAATCKLLKGIYAINFCGMMLLTCISMDRYIAIVQATKSRSLRSLTPRSKI 157
Db 161 ICLGIWALSILSLPFLFRQVFPSPNNSPVC---YEDLGHNTAKRWVLRILPHTFGFI 217
Qy 158 ICLVWGLSVLISSSTFFVNQKNTQGS-D-VCEPKYQTVS-EPIRKYQTVS-EPIRK 215
Db 218 LPLVLMFLCYGFTURTFOAHGQKRAIRVIAVLVFLCWLVPYLVLLADTLMTHTV 277
Qy 216 IPLMFIFCYTFIVKTLVQAOQSKRRKRAIRVIAVLVFLACQIPHNMLVLTAAANLGM 275
Db 278 IQETCQRNIDRALDAATEILGFLHSLNPIIYAFIGQFRNFGFLKML 325
Qy 276 NR-SCQSEKLLGYTKTVEVLAFLHCLLPVLYAFIGQFRNYFLKIL 322

RESULT 3

```

```

ENTRY A39445 #type complete
TITLE interleukin-8 receptor type A - human
ALTERNATE_NAMES interleukin-8 receptor, high-affinity
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Jan-1993 #sequence_revision 12-Apr-1996 #text_change
        29-Aug-1997
ACCESSIONS I37449; I38710; I38711; A39445
REFERENCE I37449
#authors Mollereau, C.; Muscatelli, F.; Mattei, M.G.; Vassart, G.;
        Parmentier, M.
#journal Genomics (1993) 16:248-251
#title The high-affinity interleukin 8 receptor gene (IL8RA) maps to
        the 2q33-q36 region of the human genome: Cloning of a
        pseudogene (IL8RP) for the low-affinity receptor.
#cross-references MUID:93252387
#accession I37449
#molecule_type DNA
#residues 1-350 #label RES
#cross-references EMBL:X65858; NID:g312046; PID:g312047
REFERENCE I37898
#authors Ahuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
#journal J. Biol. Chem. (1994) 269:26381-26389
#title Comparison of the genomic organization and promoter function
        for human interleukin-8 receptors A and B.
#cross-references MUID:95014476
#accession I38710
#molecule_type DNA
#residues 1-350 #label RE2
#cross-references EMBL:U11870; NID:g511804; PID:g511805
REFERENCE I38711
#molecule_type mRNA
#residues 1-16 #label RE3
#cross-references EMBL:U11871; NID:g511806; PID:g733002
REFERENCE A39445
#authors Holmes, W.E.; Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
#journal Science (1991) 253:1278-1280
#title Structure and functional expression of a human interleukin-8
        receptor.
#cross-references MUID:91368199
#accession A39445
#molecule_type mRNA
#residues 1-275, 'T', 277-350 #label HOL
#cross-references GB:M68932
GENETICS
#gene GDB:IL8RA
#cross-references GDB:135039; OMIM:146929
#map_position 2q35-2q35
#keywords G protein-coupled receptor; glycoprotein; membrane protein
SUMMARY #length 350 #molecular-weight 39791 #checksum 5406

Query Match 35.5%; Score 986; DB 2; Length 350;
Best Local Similarity 44.7%; Pred. No. 1.59e-139;
Matches 135; Conservative 75; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETTLNKYVYIYALVFLLSLGNLSVLMVLYSRGSRVTDVYLLNALAD 85
Qy 25 YCAPRRSGSSPG-YLYRIAYSILICVLGLNLIYVITFAFKKARSMTDVLNNAIAD 83
Db 86 LLFALPLTPWAASKVNG-WIFGTFCLKVSVLYKVENYFGSILLACISVDYRLAIYHATR 144
Qy 84 ILFVLTPFWAVSHATGAWFVSNAATCKLLKGIYAINFCGMMLLTCISMDRYIAIVQATK 143
Db 145 TLTKQ-RHL--VKFVCLGCGLSMNLSPFLFRQAYHPNNSPVC---YEVLGNDTAKV 198
Qy 144 SFRLSRTPRSKILCLVWGLSVLISSSTFFVNQKNTQGS-D-VCEPKYQTVS-EPIRW 201
Db 199 RMYLRILPHTFGFIVPLFVNLFCYGTFLRTLFAKMGQKRAIRVIAVLVFLCWL 258
Qy 202 KLMLGLELLFGFPFLMFEMIFCYTFIVKTLVQAOQSKRRKRAIRVIAVLVFLACQIPH 261
Db 259 NLVLLADTLMTQVIOESCERRNNIGRALDAATEILGFLHSLNPIIYAFIGQFRHGLK 318
Qy 262 NMVLLVTAANLGMNR-SCQSEKLLGYTKTVEVLAFLHCLLPVLYAFIGQFRNYFLK 320

```



```

Db 319 IL 320
QY 321 IL 322

RESULT 4
ENTRY #type complete
TITLE interleukin-8 receptor (clone 5B1a) - rabbit
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
        rabbit
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
        10-Sep-1997
ACCESSIONS A53752
REFERENCE A53752
#authors Prado, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.;
#journal Wilkenson, N.; Folco, E.; Navarro, J.
#title J. Biol. Chem. (1994) 269:12391-12394
#accession Molecular characterization of a novel rabbit interleukin-8
        receptor isotype.
#status A53752 preliminary
#molecule_type mRNA
#residues 1-358 #label PRA
#cross-references GB:L24445; NID:g437661; PID:g437662
KEYWORDS G protein-coupled receptor; transmembrane protein
SUMMARY #length 358 #molecular-weight 40631 #checksum 2000

Query Match 34.7%; Score 965; DB 2; Length 358;
Best Local Similarity 44.0%; Pred. No. 5.81e-136;
Matches 132; Conservative 80; Mismatches 77; Indels 11; Gaps 8;

Db 35 APCRSSELETSYVLLIYILVFLSLGNSLVMLVILSRSTCTSDTYLLNLAIDL 94
QY 27 APCRRSG-SSPGYRIAYSLICVLGLNLVITFAFKKARSTDTYLLNMAIDIL 85

Db 95 FATLPIWAASKVHG-WTFGTPCLKVSIVKFNVSIGILLACISDVRYLAIVHATRM 153
QY 86 FVLTPFWAVSHATGAWFSNATCKLLGIYAINFCMGLLTCTISMDRYIAIVQATKSF 145

Db 154 -IQKRHL--VFICLSMVGSLILSLPILLFNAIFPPNSPVC---YEDMGNSAKWRM 207
QY 146 RLRSRTLPKSLICLVVGLSVIISSIFVF-NQXNTGSDVCEPKYQTVSEPI-RWKL 203

Db 208 VLRILPQTFGLPLVLMFCVFTLRLFOAHMGOKHMRVFAVFLFLCWLPYNL 267
QY 204 LMLGLELLFGFFPLMFIFCFYFTVKTILVQAQNSKRHRKAIRVIAVFLVFLACQIPHN 263

Db 268 VLLTDTLMTTHVIOETCERNDDIDRALDATEILGLHSLNPIIYAFIGOKFRYGLKIL 327
QY 264 VLLVTAAANLGMNR-SCQSEKLGITKTVTEVLAFLHCLNPLVYAFIGOKFRNYFLKIL 322

RESULT 5
ENTRY #type complete
TITLE lymphocyte-specific G-protein-coupled receptor EB11 - human
ALTERNATE_NAMES Burkitt's lymphoma receptor 2; Epstein-Barr virus induced
        protein 1
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change
        10-Sep-1997
ACCESSIONS B55735
REFERENCE B55735; S52443
#authors Schweickart, V.L.; Raport, C.J.; Godiska, R.; Byers, M.G.;
#journal Eddy Jr., R.L.; Shows, T.B.; Gray, P.W.
#title Genomics (1994) 23:643-650
#accession Cloning of human and mouse EB11, a lymphoid-specific
        G-protein-coupled receptor encoded on human chromosome
        17q12-q21.2.
#status B55735 preliminary
#molecule_type mRNA
#residues 1-378 #label SCH

Db 60 FL-PLMYSVTCFVGLGNGLVILTYIYFKRLKMTDTYLLNLAVIDILFLILPFWAYS 118
QY 38 YLYRIAYSLICVLGLNLVITFAFKKARSTDTYLLNMAIDILFLVTLFPWAVSH 97

##cross-references GB:L31581; NID:g468319; PID:g468320
REFERENCE S52443
#authors Burgstahler, R.; Kempkes, B.; Staube, K.; Lipp, M.
#submission submitted to the EMBL Data Library, February 1995
#description The expression of the chemokine receptor BLR2/EB11 is
        specifically transactivated by Epstein-Barr virus nuclear
        antigen 2.
#accession S52443
#status preliminary
#molecule_type DNA
#residues 21-378 #label BUR
#cross-references EMBL:X84702
GENETICS
#gene GDB:EB11
#cross-references GDB:342065
#map_position 17q12-17q21.2
KEYWORDS G protein-coupled receptor
SUMMARY #length 378 #molecular-weight 42874 #checksum 2503

Query Match 34.6%; Score 961; DB 2; Length 378;
Best Local Similarity 45.7%; Pred. No. 2.77e-135;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;

Db 60 FL-PIMYSIICFVGLGNGLVILTYIYFKRLKMTDTYLLNLAVIDILFLILPFWAYS- 117
QY 38 YLYRIAYSLICVLGLNLVITFAFKKARSTDTYLLNMAIDILFLVTLFPWAVSH 97

Db 118 AAKSWFVGHFCKLIFAIYKMSFFSGMLLLCISIDRYVAIVQAVSAHRHRARVLLSKL 177
QY 98 ATGAWFVSNTACKLLGIYAINFCMGLLTCTISMDRYIAIVQATKSFRLSRTPRSKI 157

Db 178 SCVGIWILAVLSPELILYSD-LQRSSEQ-AMRCSLITERVE-AFTIQVQNVIGFLV 234
QY 158 ICLVWGLSVIISSTFVNQKYNTOGSDVCEPKYQTVSEPIRWKLLMLGL-ELLGFFI 216

Db 235 PLAMSCYFVILRTLQARNFERNKAIVIAVVVFIYFQLPYNGVGLAQTVANENIT 294
QY 217 PLMFIFCYFTIVKTLVQAQNSKRHRKAIRVIAVVLVFLACQIPHNVLV-VAANLGM 275

Db 295 SSTCELSQNLNIAVDVTSIACVRCVNPFLYAFIGVFKRNDLFKFKDLGCL 347
QY 276 NRSCQSEKLGITKTVTEVLAFLHCLNPLVYAFIGOKFRNYFLKILKDLWCV 328

RESULT 6
ENTRY #type complete
TITLE G protein-coupled receptor EB11 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change
        10-Sep-1997
ACCESSIONS A55735
REFERENCE A55735
#authors Schweickart, V.L.; Raport, C.J.; Godiska, R.; Byers, M.G.;
#journal Eddy Jr., R.L.; Shows, T.B.; Gray, P.W.
#title Genomics (1994) 23:643-650
#accession Cloning of human and mouse EB11, a lymphoid-specific
        G-protein-coupled receptor encoded on human chromosome
        17q12-q21.2.
#status A55735 preliminary
#molecule_type mRNA
#residues 1-378 #label SCH
#cross-references GB:L31580; NID:g468340; PID:g468341
KEYWORDS G protein-coupled receptor
SUMMARY #length 378 #molecular-weight 42941 #checksum 3827

Query Match 34.5%; Score 959; DB 2; Length 378;
Best Local Similarity 45.7%; Pred. No. 6.06e-135;
Matches 134; Conservative 74; Mismatches 78; Indels 7; Gaps 6;

Db 60 FL-PLMYSVTCFVGLGNGLVILTYIYFKRLKMTDTYLLNLAVIDILFLILPFWAYS 118
QY 38 YLYRIAYSLICVLGLNLVITFAFKKARSTDTYLLNMAIDILFLVTLFPWAVSH 97

```



```

Db 119 AK-SWIFGVYLCRGIGIYKLSFSGMLLLCISIDRYVAIQAVSRHRHARVLLISK 177
QY 98 ATGAWFSNATCKLLKGIYAINFCGMLLTCTISMDRYIAIQATKGRFLRSRTLPRSKI 157

Db 178 SCVGIWMLALFLSIPPELLYSGLOKNGEDT--LRCSLVSAQVE-ALITIOVAQMVFGFLV 234
QY 158 ICLVWGLSVLISSVFFVFNQKYNQSDVCEPKYQVSPPIRWKLLMGL-ELLFGFFI 216

Db 235 PLAMSFCLYLIIRITLQARNFERNKAIRVIIAVVVVFIYFQPYNGVLAQTVANFNIT 294
QY 217 PLMFMEFCYTFIVKTLVQAQNSKRHKRAIRVIIAVVFLACQIPHNMVLLV-TAANLGM 275

Db 295 NSCETSKOLNIADYDTYSLASVRCVNPFLYAFIGVFKFSDLKFLKGLGCL 347
QY 276 NRSCQSEKLIGYTKTVEFLAFLECLNPNLYAFIGOKFRNYFLKILKDLWCV 328

RESULT 7
ENTRY A53611 #type complete
TITLE Interleukin-8 receptor type B - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-Oct-1994 #sequence_revision 12-Apr-1996 #text_change
29-Aug-1997
ACCESSIONS 137898; 138712; A53611; A39446
REFERENCE 137898
#authors Anuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
#journal J. Biol. Chem. (1994) 269:26381-26389
#title Comparison of the genomic organization and promoter function
for human interleukin-8 receptors A and B.
#cross-references MUID:95014476
#accession 137898
#status preliminary
#molecule_type DNA
#residues 1-360 #label RES
#cross-references EMBL:U11869; NID:G511801; PID:G511803
#accession 138712
#status preliminary
#molecule_type mRNA
#residues 1-15 #label RE2
#cross-references EMBL:U11872; NID:G511808; PID:G511809; EMBL:U11873;
NID:G511810; PID:G511811; EMBL:U11874; NID:G511812;
PID:G511813; EMBL:U11875; NID:G511814; PID:G511815;
EMBL:U11876; NID:G511816; PID:G511817; EMBL:U11877;
NID:G511818; PID:G511819; EMBL:U11878; NID:G511820;
PID:G511821
REFERENCE A53611
#authors Sprenger, H.; Lloyd, A.R.; Lautens, L.L.; Bonner, T.I.;
Kelvin, D.J.
#journal J. Biol. Chem. (1994) 269:11065-11072
#title Structure, genomic organization, and expression of the human
interleukin-8 receptor B gene.
#accession A53611
#status preliminary
#molecule_type DNA
#residues 6-360 #label SPR
#cross-references GB:M99412; GB:L19593
REFERENCE A39446
#authors Murphy, P.M.; Tiffany, H.L.
#journal Science (1991) 253:1280-1283
#title Cloning of complementary DNA encoding a functional human
interleukin-8 receptor.
#cross-references MUID:91368200
#accession A39446
#status preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 6-360 #label MUR
#cross-references GB:M73969
COMMENT This receptor, unlike IL8RA, binds several peptides besides
interleukin-8, including GRO, NAP-2, and ENA-78.
GENETICS
#gene GDB:IL8RB; IL8RA
#cross-references GDB:L127868; OMIM:146928

```

```

#map_position 2q35-2q35
KEYWORDS G protein-coupled receptor; transmembrane protein
SUMMARY #length 360 #molecular_weight 40759 #checksum 3062

Query Match 34.2%; Score 951; DB 2; Length 360;
Best Local Similarity 43.8%; Pred. No. 138e-133;
Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;

Db 49 YFVVIYALVLLSLGNSLVMLVILYISRGVSRVTDVYLLNLALADLFLALFLPWAASK 108
QY 38 YLYRIAYSLLICVLGULGNLVITFAFYKARSMTDVTLLNMAIDILFVLTPFWAVSH 97

Db 109 VNG-WIEFTFFCKVYSLKKEVNFYSGILLACISVDRIYAIYHATRTLTOK-RYL--VKF 164
QY 98 ATGAWFSNATCKLLKGIYAINFCGMLLTCTISMDRYIAIQATKGRFLRSRTLPRSKI 157

Db 165 ICLSIWGLSLLALPVLFLFRRTVYSSNVSPAC--YEDMGNNNTANRWMLRLRLPQSGFI 221
QY 158 ICLVWGLSVLISSVFFVFNQK-YNTQSDVCEPKYQVSE-PIRWKLLMGLLELFGFF 215

Db 222 VPLLMLFCYGTTLRTLFKAHMGQKRAMRVIFAVVFLFLCWLPLNVLVLLADTLMTQV 281
QY 216 IPLMFMEFCYTFIVKTLVQAQNSKRHKRAIRVIIAVVFLACQIPHNMVLLVTAANLGM 275

Db 282 IOETCERNHHDRAIDALDATEILGILHSLCNPLIYAFIGOKFRHGLKIL 329
QY 276 NR-SCQSEKLIGYTKTVEFLAFLECLNPNLYAFIGOKFRNYFLKIL 322

RESULT 8
ENTRY A57160 #type complete
TITLE Chemokine (C-C) receptor 4 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change
15-Aug-1997
ACCESSIONS A57160
REFERENCE A57160
#authors Power, C.A.; Meyer, A.; Nemeth, K.; Bacon, K.B.; Hoogewerf,
A.J.; Proudfoot, A.E.I.; Wells, T.N.C.
#journal J. Biol. Chem. (1995) 270:19495-19500
#title Molecular cloning and functional expression of a novel CC
chemokine receptor cDNA from a human basophilic cell line.
#accession A57160
#status preliminary; not compared with conceptual translation
#molecule_type mRNA
#residues 1-360 #label POW
#cross-references GB:X85740
#note source clone K5-5
GENETICS
#gene GDB:CMKBR4
#cross-references GDB:677463
#map_position 3p21-3p21
KEYWORDS G protein-coupled receptor; glycoprotein; phosphoprotein;
transmembrane protein

FEATURE
40-65 #domain transmembrane #status predicted #label TM1\
76-97 #domain transmembrane #status predicted #label TM2\
112-133 #domain transmembrane #status predicted #label TM3\
151-175 #domain transmembrane #status predicted #label TM4\
208-226 #domain transmembrane #status predicted #label TM5\
243-264 #domain transmembrane #status predicted #label TM6\
291-308 #domain transmembrane #status predicted #label TM7\
29-276,110-187 #disulfide_bonds #status predicted\
72,350 #binding_site phosphate (Ser) (covalent) (by casein
kinase II) #status predicted\
145 #binding_site phosphate (Ser) (covalent) (by protein
kinase C) #status predicted\
183,194 #binding_site carbohydrate (Asn) (covalent) #status
predicted\
321 #binding_site phosphate (Thr) (covalent) (by protein
kinase C) #status predicted\
SUMMARY #length 360 #molecular_weight 41402 #checksum 4713

```



Query Match	33.48;	Score	927;	DB 2;	Length	360;			
Best Local Similarity	42.98;	Pred. No.	1.60e-129;						
Matches	129;	Conservative	81;	Mismatches	78;	Indels	13;	Gaps	12;
Db	28	PCRKEGKAPGFLPPLYSIVFGLGNSVVVLV-FKYKRLRSMTDYLNLMAISDL	86						
QY	28	PCRSGSSP-GYLYRTA-YSLICVLGLGLNLVITAF-YKKARSMTDYLNLMAIADI	84						
Db	87	LFVFSLPFGY-YAADOWFGLGLCKRMISWYLVGFYSGIFFVLMISIDRYLAIVHAV--	143						
QY	85	LFVTLFPWAVSHATGAWFSNATCKLLKGIYAINFCNGMLLTICISMDRYIAIVQATKS	144						
Db	144	FSLRARTLYGVITSLATWSAVFASLPGLFSTCYTERNHTYCKTY-SLNSTT-WKVL	201						
QY	145	FLRSLRTPRSKIICLVWGLSVIISSTVFVNQYNTQSDVCEPKYQVSPPIRWKLL	204						
Db	202	S-SLEINILGLVPLGIMLCYSMIITLQHCNEKKNAKVMFAVVVFLGLFWTPYNI	260						
QY	205	MLGLEL-LFQFFPLPMFICFYFIVKTLVQAQNSKKHKAIRVIAVVLFLACQIPHNH	263						
Db	261	VLFLTELVELEVL-QDCTFERYLDYAIQATETLAFVHCCNLPIIYFFLGKFKRYILQLF	319						
QY	264	VL-LVTAANLGMKNSQCSEKLIGYTKTVTEVLAFLHCLNPLVYAFIGQKFRNYFLKIL	322						
Db	320	K 320							
QY	323	K 323							
RESULT	9								
ENTRY	A45680	#type complete							
TITLE	G protein-coupled peptide receptor EBI 1 - human								
ORGANISM	#formal_name Homo sapiens #common_name man								
DATE	21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change								
ACCESSIONS	A45680								
REFERENCE	A45680								
#authors	Birkenbach, M.; Josefsen, K.; Yalamanchili, R.; Lenoir, G.; Kieff, E.								
#journal	J. Virol. (1993) 67:2209-2220								
#title	Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled peptide receptors.								
#cross-references	MUID:93188173								
#accession	A45680								
##status	preliminary								
##molecule_type	nucleic acid								
##residues	1-378 #label BIR								
##cross-references	GB:L08176; NID:g183484; PID:g183485								
##experimental_source	B-lymphocytes								
##note	sequence extracted from NCBI backbone (NCBIN:127094, NCBI:P127095)								
KEYWORDS	G protein-coupled receptor; transmembrane protein								
SUMMARY	#length 378 #molecular-weight 4732 #checksum 2193								
Query Match	33.48;	Score	928;	DB 2;	Length	378;			
Best Local Similarity	45.48;	Pred. No.	1.08e-129;						
Matches	133;	Conservative	76;	Mismatches	77;	Indels	7;	Gaps	7;
Db	60	FL-PIMYSIICFVLLGNLWLVITYYFKRLKMTDYLNLAVADILFLTLTLPFWAYS-117							
QY	38	YLRYIAVSLVCLGLNLVITAFYKARSMTDYLNLMAIADILFLTLTLPFWAVSH	97						
Db	118	AAKSWFGVHFCFLIPAIYKMSFFSGMLLCTISIDRYAIVQAVSHRRAVLLISKL	177						
QY	98	ATGAWFSNATCKLLKGIYAINFCNGMLLTICISMDRYIAIVQATKSFRLRSTLPSKI	157						
Db	178	SCVGSAILATVLSIPELLYSD-LQRSSEQ-AMRCSLITEHVE-AFTIQVQAWIGFLV	234						
QY	158	ICLVWGLSVIISSTVFVNQYNTQSDVCEPKYQVSPPIRWKLLMLGL-ELLGFFI	216						
Db	235	PLAMSFYLIIVITLQARNFERNKAIVIAVVVVFIVQIPYNGVLAQVAFNIT	294						



```

#title      The N terminus of interleukin-8 (IL-8) receptor confers high
#            affinity binding to human IL-8.
#cross-references MUID:94308043
#accession  A53677
#status     nucleic acid sequence not shown; not compared with
#            conceptual translation
##molecule_type DNA
##residues  1-359 ##label SUZ
##cross-references GB:I26549
##note      sequence extracted from NCBI backbone (NCBIP:149812)
REFERENCE
#authors    Lee, J.; Cacalano, G.; Camerato, T.; Toy, K.; Moore, M.W.;
#            Wood, W.I.
#journal    J. Immunol. (1995) 155:2158-2164
#title      Chemokine binding and activities mediated by the mouse IL-8
#            receptor.
#cross-references MUID:95363183
#accession  I49348
##status    preliminary; translated from GB/EMBL/DDBJ
##molecule_type DNA
##residues  1-359 ##label RES
##cross-references EMBL:U31207; NID:g950174; PID:g950175
REFERENCE
#authors    Bosic, C.R.; Gerard, N.P.; von Uexkull-Guldenband, C.;
#            Kolakowski, L.F.
#journal    J. Biol. Chem. (1994) 269:29355-29358
#title      The murine interleukin 8 type B receptor homologue and its
#            ligands. Expression and biological characterization.
#cross-references MUID:95050766
#accession  I55421
##status    preliminary; translated from GB/EMBL/DDBJ
##molecule_type DNA
##residues  1-359 ##label RE2
##cross-references GB:L13239; NID:g293665; PID:g293666
REFERENCE
#authors    Wilkie, T.M.; Chen, Y.; Gilbert, D.J.; Moore, K.J.; Yu, L.;
#            Simon, M.I.; Copeland, N.G.; Jenkins, N.A.
#journal    Genomics (1993) 18:175-184
#title      Identification, chromosomal location, and genome organization
#            of mammalian G-protein-coupled receptors.
#accession  H48909
##status    preliminary; nucleic acid sequence not shown
##molecule_type mRNA
##residues  145-258 ##label WIL
##cross-references GB:L20337; NID:g438800; PID:g438801
GENETICS
#gene       IL8rb
#status     absent
#keywords    G protein-coupled receptor; glycoprotein; transmembrane
#            protein
FEATURE
49-74      #domain transmembrane #status predicted #label TM1\
84-106     #domain transmembrane #status predicted #label TM2\
120-141    #domain transmembrane #status predicted #label TM3\
163-182    #domain transmembrane #status predicted #label TM4\
213-234    #domain transmembrane #status predicted #label TM5\
251-271    #domain transmembrane #status predicted #label TM6\
308-328    #domain transmembrane #status predicted #label TM7\
SUMMARY
#length 359 #molecular-weight 40425 #checksum 3817

Query Match      32.1%; Score 891; DB 2; Length 359;
Best Local Similarity 41.1%; Pred. No. 1,96e-123;
Matches 123; Conservative 80; Mismatches 85; Indels 11; Gaps 18;

Db      37  PCSENLNSYAVVYIVLVITLLSVGNLSWLVLYINRSCVTDVYLLNLAIDLFF 96
QY      28  PCRSG-SSPGYLYRAYSLICVLGILNVIVITPAFYKARSDTVYLLNNAIDILF 86
Db      97  ALTPYWAASKVNG-WTFGSTLCKIFSYPEVTFYSSVLLACISMDRYALVAIVHATSL- 154
QY      87  VLTLPFAVSHATGAWFNSATCKLUGIYANFNGCMULLTICISMDRYAIVQATKSR 146
Db      155  IOKRHL--VKFVCIAMWLLSVILALPILILRNPVKVNLSTLVC---YEDVGNNTSLRVV 209

```



```

RESULT      13
ENTRY       chemokine (C-C) receptor 1 - human
TITLE       C-C CKR-1; macrophage inflammatory protein-1-alpha receptor
ALTERNATE_NAMES ORGANISM   Homo sapiens #common_name man
DATE        30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
            10-Sep-1997
ACCESSIONS A45177; I55671
REFERENCE   Neote, K.; DiGregorio, D.; Mak, J.Y.; Horuk, R.; Schall, T.J.
#authors    Cell (1993) 72:415-425
#journal     Molecular cloning, functional expression, and signaling
#title       Characteristics of a C-C chemokine receptor.
#cross-references MUID:93161416
#accesion    A45177
#status      nucleic acid sequence not shown
##molecule_type mRNA
##residues   1-355 ##label NEO
##cross-references GB:I10918; NID:g292416; PID:g292417
##experimental_source HL60 cells
##note       _sequence extracted from NCBI backbone (NCBIP:124876)
REFERENCE   I55671
#authors     Gao, J.
#journal      J. Exp. Med. (1993) 177:1421-1427
#title        Structure and functional expression of the human macrophage
              inflammatory 1 alpha (MIP-lalpha)/RANTES receptor.
#cross-references MUID:93240122
#accesion    I55671
#status      preliminary; translated from GB/EMBL/DDBJ
##molecule_type mRNA
##residues   1-355 ##label RES
##cross-references GB:L10918; NID:g292416; PID:g292417
GENETICS    GDB:CMKBR1; CMKR-1
#gene         ##cross-references GDB:138446; OMIM:601159
#map_position 3p21-3p21
KEYWORDS     disulfide bond; G protein-coupled receptor; glycoprotein;
             phosphoprotein; transmembrane protein
FEATURE
36-60 domain transmembrane #status predicted #label TM1\
71-91 #domain transmembrane #status predicted #label TM2\
108-129 #domain transmembrane #status predicted #label TM3\
147-171 #domain transmembrane #status predicted #label TM4\
205-223 #domain transmembrane #status predicted #label TM5\
240-264 #domain transmembrane #status predicted #label TM6\
288-305 #domain transmembrane #status predicted #label TM7\
          #binding_site carbohydrate (Asn) (covalent) #status
          predicted\
24-273,106-183 #disulfide_bonds #status predicted\
345 #binding_site phosphate (Ser) (covalent) (by casein
           kinase II) #status predicted
SUMMARY      #length 355 #molecular_weight 41172 #checksum 7529
Query Match      30.2%; Score 839; DB 2: Length 355;
Best Local Similarity 34.9%; Pred. No. 1.16e-114;
Matches 119; Conservative 96; Mismatches 115; Indels 11; Gaps 9;
Db      22 TPCKVNERAFAQGLPPLYSLVFVGILGNLVLVLYOYRKLNKTSTYLNLNLAISDL 81
Qy      27 APCRRSGSP-GY-LYRTAYSLLICVLGCLGNLVITFYAFYKKARSWTDVYLNMAIADI 84
Db      82 LFEFLTPWIDYKLKDWDVFGDAMCKILSGFGYYTGYLSIFFILLITDRLYAIVHAV-- 139
Qy      85 LFVLTLPEAVSHATGAWFVNATCKLKLGIAINFNCGMGLLTCTCSMDRYIAVOATKS 144
Db      140 FALRAVTTFGVTTIIIIWAAILASMPGLYSFKTSQTWEHTHCTS LHFP--HSLERWKL 197
Qy      145 FLRSRTLPRSRICI LVNWGLSVIISSSTFFVENKYNTQGSVDVEPKYTQVSPIR-WKL 203
Db      198 FQ-AALKNLFGVLPLLVMICYTGIIKIILLRPNEKSKAVRLFIVMIIFLFWTPTPN 256
Qy      204 LMIGLEG-LFGGFPLEMFMTFTTYIKTVLTAQAQRKHKAIRVIIVAIVLVFLAACQIPHN 262
```



```
##status      preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues    1-359 #label RES
##cross-references EMBL:028406; NID:g881551; PID:g881552
SUMMARY      #length 359 #molecular-weight 41857 #checksum 8333

Query Match      30.0%; Score 835; DB 2; Length 359;
Best Local Similarity 38.4%; Pred.No. 5.47e-114;
Matches 108; Conservative 89; Mismatches 76; Indels 8; Gaps 7;

Db 45 YSLVFIIGLGNMNVLLILIKYRKQLQIMTWIYLFNLAIISDLLFLFTVPFW-IHYVLWNEW 103
QY 44 YSLICVLGLIGNILVITFAFYKKARSMTDVYLLNMAIADILFVLTLPFWAVSHAT-GAW 102

Db 104 GFQHYMCKMLSGFYLYALYSEIFFIILLTIDRYLAIVHAV--FALRARTVTFATITSIT 161
QY 103 VFSNATCKLLKGIYAINFGNCGMLLLTCISMDRYAIQVATKSPRLRSRTLPRSKIICLV 162

Db 162 WGLAGLAALPEEFIFHESQDSFGESCSPRYPE-GEEDSWKRFHAI RMNI-FGLALPLLVM 219
QY 163 WGLSVIISSSTFVFNQKYNTOGSDVCEPKYQTVSEPIRWK-LMLGLLELLFGFFIPLMF 221

Db 220 VICYSIIKTLRCPNKKKKAIRLIFWVMIVFFIFWTYPYNLVLLFSAFHSTFLETSC 279
QY 222 IFCTYFIVKTLVQAQNSKRKAIRVIAVVLVFLACOIPHNWVLLVTAANLGKMRSCQS 281

Db 280 SKHLDLAMOQTEVIAYTHCCVNPVIYAFVGERPKH-LRLF 319
QY 282 EKLIGYTKTVTEVLAFHCCCLNPVLYAFIGKPRNYFLKIL 322
```

Search completed: Fri Nov 13 12:04:53 1998  
Job time : 72 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

```
Run on:      Fri Nov 13 12:05:10 1998;  MasPar time 10.77 Seconds
           850.044 Million cell updates/sec
Tabular output not generated.
```

```

Title: >US-08-887-977-10
Description: (1-365) from US0887977.pep
Perfect Score: 2779
Sequence: 1 MFSPVKILQCSILHIQL.....NISRTSETNDNDNASSETM 365

```

Scoring table: PAM 150  
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

```
Database: swiss-prot35
          1:swiss1
```

Statistics: .. Mean 49.069; Variance 98.526; scale 0.498

Pred..No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result	No.	Score	Query		DB	ID	Description	Pred. No.
			Match	Length				
1	1	2445	88.0	369	1	CKR6_HUMAN	C-C CHEMOKINE RECEPTOR	0.00e+00
2	2	994	35.8	350	1	IL8A_PANTR	HIGH AFFINITY INTERLEU	1.02e-170
3	3	994	35.8	350	1	IL8A_GORGO	HIGH AFFINITY INTERLEU	1.02e-170
4	4	994	35.8	350	1	IL8A_PONPY	HIGH AFFINITY INTERLEU	1.02e-170
5	5	990	35.6	355	1	IL8A_RABIT	HIGH AFFINITY INTERLEU	6.81e-170
6	6	984	35.4	350	1	IL8A_HUMAN	HIGH AFFINITY INTERLEU	1.17e-168
7	7	978	35.2	360	1	IL8B_BOVIN	HIGH AFFINITY INTERLEU	2.02e-167
8	8	965	34.7	358	1	IL8B_RABIT	HIGH AFFINITY INTERLEU	9.57e-165
9	9	961	34.6	378	1	CKR7_HUMAN	C-C CHEMOKINE RECEPTOR	6.36e-164
10	10	959	34.5	353	1	IL8B_MACMU	HIGH AFFINITY INTERLEU	1.64e-163
11	11	959	34.5	378	1	CKR7_MOUSE	C-C CHEMOKINE RECEPTOR	1.64e-163
12	12	951	34.2	353	1	IL8B_PANTR	HIGH AFFINITY INTERLEU	7.25e-162
13	13	951	34.2	360	1	IL8B_HUMAN	HIGH AFFINITY INTERLEU	7.25e-162
14	14	945	34.0	353	1	IL8B_GORGO	HIGH AFFINITY INTERLEU	1.24e-160
15	15	927	33.4	360	1	CKR4_HUMAN	C-C CHEMOKINE RECEPTOR	6.19e-157
16	16	920	33.1	359	1	IL8B_RAT	HIGH AFFINITY INTERLEU	1.69e-155
17	17	911	32.8	349	1	IL8A_RAT	HIGH AFFINITY INTERLEU	1.19e-153
18	18	894	32.2	357	1	GC96_HUMAN	PROBABLE G PROTEIN-COU	3.64e-150
19	19	891	32.1	359	1	IL8B_MOUSE	HIGH AFFINITY INTERLEU	1.50e-149
20	20	870	31.3	360	1	CKR4_MOUSE	C-C CHEMOKINE RECEPTOR	2.99e-145
21	21	839	30.2	355	1	CKR1_HUMAN	C-C CHEMOKINE RECEPTOR	6.49e-139
22	22	835	30.0	359	1	CKR3_MOUSE	PROBABLE C-C CHEMOKINE	4.26e-138
23	23	824	29.7	368	1	CCR3_HUMAN	C-X-C CHEMOKINE RECEPT	7.48e-136











AC P21109;  
 DT 01-FEB-1991 (REL. 17, CREATED)  
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR A (IL-8R A) (CXCR-1).  
 GN IL8RA OR CXCR1.  
 OS ORYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; LAGOMORPHA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91378994.  
 RA BECKMANN M.P., MUNGER W.E., KOZLOSKY C., VANDENBOS T., PRICE V.,  
 RA LYMAN S., GERARD N.P., GERARD C., CERRETTI D.P.,  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 179:784-789(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ALBINO; TISSUE-NEUTROPHILS;  
 RX MEDLINE; 91056034.  
 RA THOMAS K.M., PYUN H.Y., NAVARRO J.,  
 RL J. BIOL. CHEM. 265:20061-20064 (1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-NEUTROPHILS;  
 RX MEDLINE; 92148149.  
 RA LEE J., KUANG W.-J., RICE G.C., WOOD W.I.;  
 RL J. IMMUNOL. 148:1261-1264 (1992).  
 CC -!- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL  
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR  
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A  
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND  
 CC MESSENGER SYSTEM.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: NEUTROPHILS.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -!- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE THE RECEPTOR FOR  
 CC FMET-LEU-PHE (N-FORMYL PEPTIDE RECEPTOR).  
 DR EMBL; M74240; G165439; -;  
 DR EMBL; M58021; G165443; -;  
 DR EMBL; M82873; G165441; -;  
 DR PIR; A23669; A23669.  
 DR PIR; A46483; A46483.  
 DR PIR; JQ1231; JQ1231.  
 DR GCRDB; GCR\_0107; -;  
 DR GCRDB; GCR\_0108; -;  
 DR GCRDB; GCR\_0298; -;  
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR; 1.  
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW CHEMOTAXIS.  
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 41 67 1 (POTENTIAL).  
 FT DOMAIN 68 73 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 74 92 2 (POTENTIAL).  
 FT DOMAIN 93 114 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 115 138 3 (POTENTIAL).  
 FT DOMAIN 139 159 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 160 184 4 (POTENTIAL).  
 FT DOMAIN 185 204 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 205 232 5 (POTENTIAL).  
 FT DOMAIN 233 247 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 248 270 6 (POTENTIAL).  
 FT DOMAIN 271 290 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 291 313 7 (POTENTIAL).  
 FT DOMAIN 314 355 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 21 21 POTENTIAL.  
 FT DISULFID 115 192 BY SIMILARITY.  
 FT CONFLICT 90 111 DFLALPMP1WAVSKRGWIFG ->  
 FT CONFLICT 146 147 PAFCPDHAYLRGKQKRLDFR (IN REF. 2).  
 FT CONFLICT 204 204 HA -> QS (IN REF. 2).  
 FT CONFLICT 287 288 R -> C (IN REF. 2).  
 FT CONFLICT 288 288 DI -> EL (IN REF. 2).  
 FT CONFLICT 355 355 AA; 40622 MW; 2EB3947D CRC32;  
 SQ SEQUENCE

Query Match 35.6%; Score 990; DB 1; Length 355;  
 Best Local Similarity 45.1%; Pred. No. 6.81e-170;  
 Matches 130; Conservative 78; Mismatches 70; Indels 10; Gaps 7;  
 DB 45 YVWVVIYALVFLSLGNSLVMLVILYSRSNRSVTDVYLLNLAMADLEFALTMPIAVSK 104  
 QY 38 YLRIAYSGLCVGLGNLIVITFAFYKARSMVDVLLNMAIDILFVLTLEFVAVSH 97  
 DB 105 EKG-WIFGTPCKVSVLKEVNFYSGILLACISVDRLAIVHATRLTQK-RHL--VKF 160  
 QY 98 ATGAWFVSATCKLLKGIYAINFCMGLLITCISMDRYAIVAQATKSFRLSRTPRSKI 157  
 DB 161 ICLGIWALSILSLPFLPQVFPNNSPVC---YEDLGHNTAKRWVRILPHITGFI 217  
 QY 158 ICLVWGLSVIISSSTFVFNQKYNQTQSD-VCEPKYQTVS-EPTRWKLMLGLELLEFGF 215  
 DB 218 LPLLVMFCYGETLRTLFQAHGQKHRAVEIVAVVLIFELLCWLPYNLVLLADFLMETHV 277  
 QY 216 IPLMFIMFCYTFIVKTLVQONSKRAIRVIAVVLVFLACQIPHNNVLLVTAANLGM 275  
 DB 278 IQETCQRNIDRALDATEILGELHSLNPIIYAFIGNFNGFLKML 325  
 QY 276 NR-SCOSEKLIQYTKVTEVLAFELHCLNPLVYAFIQKERNYFLKIL 322  
 RESULT 6  
 ID IL8A\_HUMAN STANDARD; PRT; 350 AA.  
 AC P25024;  
 DT 01-MAY-1992 (REL. 22, CREATED)  
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR A (IL-8R A) (IL-8 RECEPTOR TYPE  
 DE 1) (CXCR-1) (CDW128).  
 GN IL8RA OR CXCR1.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; PRIMATES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91368199.  
 RA HOLMES W.E., LEE J., KUANG W.-J., RICE G.C., WOOD W.I.;  
 RL SCIENCE 253:1278-1280(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX MEDLINE; 93205012.  
 RA CERRETTI D.P., KOZLOSKY C.J., VANDEN BOS T., NELSON N., GEARING D.P.,  
 RA BECKMANN M.P.;  
 RL MOL. IMMUNOL. 30:359-367(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93252387.  
 RA MOLLEREAU C., PASSAGE E., MATTEI M.-G., VASSART G., PARMENTIER M.;  
 RL GENOMICS 16:248-251(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PLACENTA;  
 RX MEDLINE; 95014476.  
 RA AHUTA S.K., SHETTY A., TIFFANY H.L., MURPHY P.M.;  
 RL J. BIOL. CHEM. 269:26381-26389(1994).  
 RN [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE; 92355587.  
 RA LEE J., HORUK R., RICE G.C., BENNETT G.L., CAMERATO T., WOOD W.I.;  
 RL J. BIOL. CHEM. 267:16283-16287(1992).  
 CC -!- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL  
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR  
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A  
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND  
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY  
 CC AND TO MGSA (GRO) WITH A LOW AFFINITY.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.











QY 98 ATGAWFVNATCKLLKGIYAINFNCGMLLTCISMDRYIAIVQATKSFRLRSRTILPRSKI 157  
 Db 178 SCVGIWILATVLSPELISD-LQSSSEQ-AMRSLITEHYE-AFIIQVQMWIGFLV 234  
 QY 158 ICLVWGLSVLISSSTFVNQKINTQGSVDCPKYQTVTSEPRLWLLMLGL-ELLFGFFI 216  
 Db 235 PLUMSCYLVIIITLLQARFERNKAIKVIIVVVFVIFVOLFPGYGVVLAQTVANFNIT 294  
 QY 217 PLMEFICYFTIVKTLVQAQNSKRKAIRVIAVVLVFLACOIPHNWLLV-TAANLGM 275  
 Db 295 SSTCELSQNLNAYDVVSYLACVRCVNPFLVYAFIVGKFRNDLFLKFDLGL 347  
 QY 276 NRSQSEKLGITKTVTEVLAFLHCLNPLVYAFIVGKFRNVLKILKDLWCV 328

RESULT 10  
 ID IL8B\_MACMU STANDARD; PRT; 353 AA.  
 AC Q28519;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).  
 GN IL8RB OR CXCR2.  
 OS MACACA MULATTA (RHESUS MACAQUE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; PRIMATES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96175151.  
 RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,  
 RA LOPEZ-LARREA C.;  
 RL IMMUNOGENETICS 43:261-267(1996).  
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL  
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR  
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A  
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND  
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY  
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL; X91116; E198176; -  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR; 1.  
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW CHEMOTAXIS.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 46 72 1 (POTENTIAL).  
 FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 82 102 2 (POTENTIAL).  
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 118 139 3 (POTENTIAL).  
 FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 161 180 4 (POTENTIAL).  
 FT DOMAIN 181 205 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 206 228 5 (POTENTIAL).  
 FT DOMAIN 229 240 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 249 270 6 (POTENTIAL).  
 FT DOMAIN 271 291 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 292 312 7 (POTENTIAL).  
 FT DOMAIN 313 >353 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 116 193 BY SIMILARITY.  
 FT CARBOHYD 19 19  
 FT NON\_TER 353 353  
 SQ SEQUENCE 353 AA; B592FE64 CRC32;

Query Match 34.5%; Score 959; DB 1; Length 353;  
 Best Local Similarity 43.0%; Pred. No. 1.64e-163;  
 Matches 129; Conservative 82; Mismatches 78; Indels 11; Gaps 8;

Db 34 APCRPESLEINKFVVIYALVFLSLGNSLVMLVILSVGRSVTDVYLLNALADLL 93  
 QY 27 APCRRSSGSPG-YLYRIAYSLICVLGLGNILVITFAFYKKARSMTDVLNMAIADIL 85

Db 94 FALTPLPIAAKNGV-WIFGTFLCKVYVLLKEVNFYSGILLACISVDYRLAIIVHATRTL 152  
 QY 86 FVTLPLFWAVSHATGAWVFNATCKLLKGIYAINFNCGMLLTCISMDRYIAIVQATKSF 145  
 Db 153 TOK-RYL--VKFICLSIWGLSLLLALPVLFFRRTVYSNVSFAC---YEDMGNTANRM 206  
 QY 146 RLRSRTLPRSKIIICLVWGLSVLISSSTFVNQK-YNTQGSVDCPKYQTVSE-PIRWKL 203  
 Db 207 LLRLIPOSFGFIVPLLMFLCYGFTLTTLFKAHQKQHRAMRVIFAVVLFLCWLDPYSL 266  
 QY 204 LMLGLELLFGFFIPLMEFICYFTIVKTLVQAQNSKRKAIRVIAVVLVFLACQIPHN 263  
 Db 267 VLLADTLMTQVIQETCERNHIDRALDATEILGILHCLNPLIYAFIVGKFRHGLKIL 326  
 QY 264 VLLVTANLGMNR-SCQSEKLGITKTVTEVLAFLHCLNPLVYAFIVGKFRNFIKIL 322

RESULT 11  
 ID CKR7\_MOUSE STANDARD; PRT; 378 AA.  
 AC P47774;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE C-C CHEMOKINE RECEPTOR TYPE 7 PRECURSOR (C-C CKR-7) (CCR7) (EBV-  
 DE INDUCED G PROTEIN-COUPLED RECEPTOR 1) (EB1).  
 GN CKR7 OR EB1 OR EB1H.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B6/CBA; TISSUE=THYMUS;  
 RX MEDLINE; 95154835.  
 RA SCHWICKART V.L., RAPPORT C.J., GODISKA R., BYERS M.G., EDDY R.L. JR.,  
 RA SHOWS T.B., GRAY P.W.;  
 RL GENOMICS 23:643-650(1994).  
 CC -1- FUNCTION: RECEPTOR FOR THE ELC/MIP3B CHEMOKINE. PROBABLE MEDIATOR  
 CC OF EBV EFFECTS ON B LYMPHOCYTES OR OF NORMAL LYMPHOCYTE FUNCTIONS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL; L31580; G468341; -  
 DR MGD; MG1:103011; CKR7.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR; 1.  
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.  
 FT SIGNAL 1 24  
 FT CHAIN 25 378 C-C CHEMOKINE RECEPTOR TYPE 7.  
 FT DOMAIN 25 59 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 60 86 1 (POTENTIAL).  
 FT DOMAIN 87 95 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 96 116 2 (POTENTIAL).  
 FT DOMAIN 117 130 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 131 152 3 (POTENTIAL).  
 FT DOMAIN 153 170 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 171 191 4 (POTENTIAL).  
 FT DOMAIN 192 219 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 220 247 5 (POTENTIAL).  
 FT DOMAIN 248 263 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 264 289 6 (POTENTIAL).  
 FT DOMAIN 290 313 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 314 331 7 (POTENTIAL).  
 FT DOMAIN 332 378 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 36 36  
 FT DISULFID 129 210 BY SIMILARITY.  
 SQ SEQUENCE 378 AA; 42941 MW; 9602A43B CRC32;

Query Match 34.5%; Score 959; DB 1; Length 378;  
 Best Local Similarity 45.7%; Pred. No. 1.64e-163;  
 Matches 134; Conservative 74; Mismatches 78; Indels 7; Gaps 6;

Db 60 FL-PLMNVICFVGLGNGLVILTYIFKRLKMTDYLNLAVADILFLILPFWAYSE 118  
 QY 38 YLYRIAYSLICVLGLGNILVITFAFYKKARSMTDVLNMAIADILFLILPFWAVSH 97



38 YLYRIAYSLICVLGLLGNILVITFYFKYKARSMTDVLINMAIDOLFVLTLPLFPWAVSH 97

106 VNG-WIEGTFLCKVSVLLKEVNEYSGLILLACISVDRLAIIVHATRLTQK-RYL--VKF 161

98 ATGAWFSNATCKLLGIYAINFCNGLLUTCLISMDRYIAIVQATSKFSLRSTLPRSKI 157

162 ICLISLISLALLPALPLFRRTYSSNVSPAC---YEDMGNTANWRMLRLMLPQSGFGPI 218

158 ICLVWGLSVIISSIFVFNQK-YNTQSDVCEPKQYVSE-PIRWKLLMLGLELLFGFE 215

219 VPLLIMFCVGTFLRLTFLKAHQGKHRAKRVIPAVVLIIFLLCWLPVYNVLVLLADTLMTQV 278

216 IPLMFICFTFVTKTLVQAQNSKRKAIRVIAVTVLVFLACQIPHNVLVLTAAANLGM 275

279 IQETCERNHINRDLATETILGLTHSLCNPLIYAFIGOKERHGLLKLIL 326

276 NR-SCOSEKLEIGTKVTVEIAFLHCLLNPLVYAFIGOKFRNYFLKIL 322

RESULT 13

ID IL8B\_HUMAN STANDARD; PRT; 360 AA.

AC P25025;

DT 01-MAY-1992 (REL. 22, CREATED)

DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MGSA RECEPTOR) (IL-8 RECEPTOR TYPE 2).

GN IL8RB OR CXCR2.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

[1]

RX SEQUENCE FROM N.A.

RP MEDLINE; 91368200.

RA MORPHY P.M., TIFFANY H.L.;

RA SCIENCE 253:1280-1283(1991).

[2]

RN SEQUENCE FROM N.A., AND CHARACTERIZATION.

RP MEDLINE; 93205012.

RX CERRETTI D.P., KOZLOSKY C.J., VANDEN BOS T., NELSON N., GEARING D.P.,

RA SPRENGER H., LLOYD A.R., LAUTENS L.L., BONNER T.I., KELVIN D.J.;

RA J. BIOL. CHEM. 269:11065-11072(1994).

[4]

RN SEQUENCE FROM N.A.

RP TISSUE-PLACENTA.

RX MEDLINE; 95014476.

RA AHUJA S.K., SHETTY A., TIFFANY H.L., MURPHY P.M.;

RA J. BIOL. CHEM. 269:26381-26389(1994).

[5]

RN CHARACTERIZATION.

RP MEDLINE; 92355587.

RA LEE J., HORUK R., RICE G.C., BENNETT G.L., CAMERATO T., WOOD W.I.;

J. BIOL. CHEM. 267:16283-16287(1992).

CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL

CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR

CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A

CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND

CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY

CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

DR EMBL; M73969; G1109691; -

DR EMBL; M94582; G186378; -

DR EMBL; M98412; G576679; -

DR EMBL; L19593; G559054; -

DR EMBL; U11869; G511803; -

DR FIC; A39446; A39446; -

DR FIC; A53611; A53611; -



DR GCRDB: GCR\_0077: -  
 DR GCRDB: GCR\_0610: -  
 DR GCRDB: GCR\_1001: -  
 DR MM: 146928: -  
 DR PROSITE: PS00237: G\_PROTEIN\_RECEPTOR; 1.  
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW CHEMOTAXIS.  
 FT DOMAIN 1 48 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 49 75 1 (POTENTIAL).  
 FT DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 85 105 2 (POTENTIAL).  
 FT DOMAIN 106 120 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 121 142 3 (POTENTIAL).  
 FT DOMAIN 143 163 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 164 183 4 (POTENTIAL).  
 FT DOMAIN 184 208 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 209 231 5 (POTENTIAL).  
 FT DOMAIN 232 251 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 252 273 6 (POTENTIAL).  
 FT DOMAIN 274 294 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 295 315 7 (POTENTIAL).  
 FT DOMAIN 316 360 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 22 22 POTENTIAL.  
 FT DISULFID 119 196 BY SIMILARITY.  
 SQ SEQUENCE 360 AA; 40759 MW; 135CEAE CRC32;

Query Match 34.2%; Score 951; DB 1; Length 360;  
 Best Local Similarity 43.8%; Pred. No. 7.25e-162;  
 Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;  
 Db 49 YFVVIIVALLVLLSLGSLVNLVLYSRVGRSVTDVYLLNLALADLLFALTLPWAASK 108  
 QY 38 YLYRIAYSLICVLGGLNVLVITFAFKKARSDVYLLNNAIADILFVLTLPFWAVSH 97  
 Db 109 VNG-WIEFTLCKVSVLLKEVNFYSGILLACISVDRIYLAIVHATRLTQK-RYL--VKF 164  
 QY 98 ATGAWFVSNAATKLLGIYAINFNGMLLTLCISMDRYIAIVQATKSPRLSRTPRSKI 157  
 Db 165 ICLSWGLSLLALPVLLFRRTVPSNVSPAC---YEDMGNTANWRMLRLPQSGFTI 221  
 QY 158 ICLVWGLSVIISSTFVFNOK-YNTQSDVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215  
 Db 222 VPLIMLFCYFTLRTLFKAHMGOKHRMRYIFAVLFLCWLPLYNVLADTLMTQV 281  
 QY 216 IPLMFIFCYFTFVKTQVQNSKHKRAIRVIAVVLVFLACQIPHNVLVTAANLGM 275  
 Db 282 IQETCERNHIDRALDATEILGILHSCNLPLIYAFIQGKFRGLLKIL 329  
 QY 276 NR-SCQSEKLGITVTVTEVLAFLHCLLPVLYAFIQGKFRNYFLKIL 322

RESULT 14  
 ID IL8B GORGO STANDARD; PRT; 353 AA.  
 AC Q28422;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).  
 GN IL8RB OR CXCR2.  
 OS GORILLA GORILLA GORILLA (LOWLAND GORILLA).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; PRIMATES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96175151.  
 RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,  
 RA LOPEZ-LARREA C.;  
 RL IMMUNOGENETICS 43:261-267(1996).  
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL  
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR  
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A  
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND  
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY.

CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL: X91114; E198175; -  
 KW PROSITE: PS00237; G\_PROTEIN\_RECEPTOR; 1.  
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW CHEMOTAXIS.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 46 72 1 (POTENTIAL).  
 FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 82 102 2 (POTENTIAL).  
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 118 139 3 (POTENTIAL).  
 FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 161 180 4 (POTENTIAL).  
 FT DOMAIN 181 205 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 206 228 5 (POTENTIAL).  
 FT DOMAIN 229 248 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 249 270 6 (POTENTIAL).  
 FT DOMAIN 271 291 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 292 312 7 (POTENTIAL).  
 FT DOMAIN 313 >353 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 116 193 BY SIMILARITY.  
 FT CARBOHYD 19 19 POTENTIAL.  
 FT NON\_TER 353 353  
 SQ SEQUENCE 353 AA; 39919 MW; 4AF43313 CRC32;

Query Match 34.0%; Score 945; DB 1; Length 353;  
 Best Local Similarity 43.8%; Pred. No. 1.24e-160;  
 Matches 126; Conservative 79; Mismatches 73; Indels 10; Gaps 7;  
 Db 46 YFVVIIVALLVLLSLGSLVNLVLYSRVGRSVTDVYLLNLALADLLFALTLPWAASK 105  
 QY 38 YLYRIAYSLICVLGGLNVLVITFAFKKARSDVYLLNNAIADILFVLTLPFWAVSH 97  
 Db 106 VNG-WIEFTLCKVSVLLKEVNFYSGILLACISVDRIYLAIVHATRLTQK-RYL--VKF 161  
 QY 98 ATGAWFVSNAATKLLGIYAINFNGMLLTLCISMDRYIAIVQATKSPRLSRTPRSKI 157  
 Db 162 ICLSWGLSLLALPVLLFRRTVPSNVSPVC---YEDMGNTANWRMLRLPQSGFTI 218  
 QY 158 ICLVWGLSVIISSTFVFNOK-YNTQSDVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215  
 Db 219 VPLIMLFCYFTLRTLFKAHMGOKHRMRYIFAVLFLCWLPLYNVLADTLMTQV 278  
 QY 216 IPLMFIFCYFTFVKTQVQNSKHKRAIRVIAVVLVFLACQIPHNVLVTAANLGM 275  
 Db 279 IQETCERNHINQALDATEILGILHSCNLPLIYAFIQGKFRGLLKIL 326  
 QY 276 NR-SCQSEKLGITVTVTEVLAFLHCLLPVLYAFIQGKFRNYFLKIL 322

RESULT 15  
 ID CKR4 HUMAN STANDARD; PRT; 360 AA.  
 AC P51679;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE C-C CHEMOKINE RECEPTOR TYPE 4 (C-C CKR-4) (CXCR-4) (K5-5).  
 GN CKMR4.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; PRIMATES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95370289.  
 RA POWELL C.A., MEYER A., NEMETH K., BACON K.B., HOOGWERF A.J.,  
 RA PROUDFOOT A.E.I., WELLS T.N.C.;  
 RL J. BIOL. CHEM. 270:19495-19500(1995).  
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA  
 CC AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY INCREASING THE  
 CC INTRACELLULAR CALCIUM IONS LEVEL.



```
CC      -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC      -!- TISSUE SPECIFICITY: RESTRICTED TO LEUKOCYTE-RICH TISSUES.
CC      -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR      EMBL; X85740; G971452; -
DR      PROSITE; PS00237; G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
KW      G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT      DOMAIN 1 39 67 1 (POTENTIAL).
FT      DOMAIN 2 68 77 1 (POTENTIAL).
FT      DOMAIN 3 78 98 2 (POTENTIAL).
FT      DOMAIN 4 99 111 2 (POTENTIAL).
FT      DOMAIN 5 112 133 3 (POTENTIAL).
FT      DOMAIN 6 134 150 3 (POTENTIAL).
FT      DOMAIN 7 151 175 4 (POTENTIAL).
FT      DOMAIN 8 176 206 5 (POTENTIAL).
FT      DOMAIN 9 207 226 5 (POTENTIAL).
FT      DOMAIN 10 227 242 6 (POTENTIAL).
FT      DOMAIN 11 243 267 6 (POTENTIAL).
FT      DOMAIN 12 268 284 7 (POTENTIAL).
FT      DOMAIN 13 285 308 7 (POTENTIAL).
FT      DOMAIN 14 309 360 7 (POTENTIAL).
FT      CARBOHYD 183 183 183 POTENTIAL.
FT      CARBOHYD 194 194 194 POTENTIAL.
FT      DISULFID 110 187 BY SIMILARITY.
SQ      SEQUENCE 360 AA; 41402 MW; 8738E75E CRC32;

Query Match      33.4%; Score 927; DB 1; Length 360;
Best Local Similarity 42.9%; Pred.No. 6.19e-157;
Matches 129; Conservative 81; Mismatches 78; Indels 13; Gaps 12;

Db      28 PCTGKIGAFGLFPLPLSLVFLVFGLLGNSVVVLV-FYKRLRSMTDVYLLNLAISDL 86
QY      || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
28 PCRRSGSSP-GVLYRIA-YSLICVLGLGLNLVYITFAF-YKARSMTDVYLLNNAIADI 84
||| : |||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      87 LFVFSLPFNGY-YAADOVYFGLGLCKMISWMLVGYSGIFFVFMMSIDRYLAIVHAV-- 143
||| : |||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      85 LFVILPLFWAVSHATGAWFSNATCKLLKGIYAINFCGMLLLCISMDRYIAIVQATKS 144
||| : |||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      144 FSLRARTLYGVITSLATWSVAVFASLPGLFSTCYTERNHYCKTRY-SLNSTT-WKVL 201
| : ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      145 FLRSRTLPKSKICLVVWGLSVIISSTFVNQKYNQSGDVCEPKYQVSEPIRWKLL 204
||| : |||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      202 S-SLEINILGLVPIGIMLFCYSMIINTLOHCKNEKNKAVKMFVAVVFLGFWTPYNI 260
: || : : || : || : || : || : || : || : || : || : || : || : || : || : ||
QY      205 MLGLEL-LFGFFIPLFMFICFYTFIVKTLVQAQNSKRKAIRVIAVVLVLAQIQPHNM 263
||| : |||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      261 VLFELETLVEVL-QDCTERYLDYAQAETLAFVHCCLNPIIYFFLGEKFKYILQLF 319
|| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      264 VL-LVTAANLGRMNRSCQEKLGITRTVTPEVLAFHCCLNPLVIAFGQKFRNYFLKIL 322
|| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      320 K 320
QY      323 K 323

Search completed: Fri Nov 13 12:05:50 1998
Job time : 40 secs.
```



\*\*\*\*\*

WIREH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Nov 13 12:06:08 1998; MasPar time 21.91 Seconds  
Tabular output not generated. 829.627 Million cell updates/sec

Title: >US-08-887-977-10  
Description: (1-365) from US08887977.ppt  
Perfect Score: 2779  
Sequence: 1 MFSTPVKILLCQSILHITQL.....NISQTSATDNDNASSFTM 365

Scoring table:  
PAM 150  
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: spiremb16  
1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 47.707; Variance 110.430; scale 0.432

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2051	73.8	367	11	054689 G PROTEIN-COUPLED RECE	0.00e+00
2	833	30.0	368	13	042444 INTERLEUKIN-8-LIKE REC	1.60e-119
3	824	29.7	359	11	055169 RECEPTOR PROTEIN CCR3	6.32e-118
4	820	29.5	415	4	015185 G PROTEIN-COUPLED RECE	3.24e-117
5	817	29.4	359	11	054814 CHEMOKINE RECEPTOR CCR	1.10e-116
6	799	28.8	352	6	062743 CHEMOKINE RECEPTOR CCR	1.70e-113
7	799	28.8	352	6	062746 CHEMOKINE RECEPTOR CCR	1.70e-113
8	798	28.7	352	6	062745 CHEMOKINE RECEPTOR CCR	2.55e-113
9	796	28.6	352	6	062744 CHEMOKINE RECEPTOR CCR	5.76e-113
10	786	28.3	352	6	018770 CCR5 RECEPTOR (FRAGMEN	3.39e-111
11	776	27.9	352	4	015538 CCR5 RECEPTOR (FRAGMEN	1.98e-109
12	772	27.8	373	11	055193 CHEMOKINE RECEPTOR CCR	1.01e-108
13	769	27.7	352	6	018772 CCR5 RECEPTOR (FRAGMEN	3.42e-108
14	769	27.7	352	6	018771 CCR5 RECEPTOR (FRAGMEN	3.42e-108
15	766	27.6	384	4	000537 CHEMOKINE RECEPTOR CCR	1.16e-107
16	763	27.5	384	4	000590 CG-CHEMOKINE RECEPTOR	3.92e-107
17	756	27.2	360	6	018793 CHEMOKINE RECEPTOR	6.73e-106
18	739	26.6	333	4	014694 CCR5 RECEPTOR (FRAGMEN	6.67e-103
19	734	26.4	373	11	008707 CHEMOKINE (C-C) RECEPT	5.07e-102
20	722	26.0	382	11	009027 CCR10-RELATED RECEPTOR	6.55e-100

21	717	25.8	360	4	060835 CXCR4 GENE ENCODING RE	4.36e-99
22	710	25.5	352	6	046428 ALPHA-CHEMOKINE RECEPT	8.43e-98
23	708	25.5	352	6	082747 CHEMOKINE RECEPTOR CX	1.89e-97
24	683	24.6	383	14	Q89609 G PROTEIN-COUPLED RECE	4.60e-93
25	513	18.5	353	13	P79960 MESENCHYME-ASSOCIATED	1.35e-63
26	506	18.2	372	11	O70526 BRADYKININ B2 RECEPTOR	2.13e-62
27	482	17.3	332	11	O62973 CHEMOKINE RECEPTOR LCR	2.62e-58
28	482	17.3	370	4	O15132 P2Y5-LIKE RECEPTOR	2.62e-58
29	460	16.6	383	13	O42324 MU-OPIOID RECEPTOR	1.41e-54
30	457	16.4	238	13	Q92158 ANGIOTENSIN II RECEPTO	4.55e-54
31	449	16.2	185	13	O42445 CX3 CHEMOKINE RECEPTOR	1.02e-52
32	441	15.9	381	13	O42402 ORNITHOKININ RECEPTOR	2.28e-51
33	438	15.8	361	11	O35811 G-PROTEIN COUPLED RECE	7.30e-51
34	437	15.7	373	13	O57585 OPIOID RECEPTOR HOMOLO	1.08e-50
35	436	15.7	404	4	O35218 G-PROTEIN COUPLED RECE	1.58e-50
36	421	15.1	395	11	O84166 ADRENOMEDULLIN RECEPTO	5.22e-48
37	415	14.9	303	13	P70058 ANGIOTENSIN RECEPTOR R	5.27e-47
38	414	14.9	374	13	O57466 G PROTEIN COUPLED P2Y	7.75e-47
39	394	14.2	344	4	O00421 CCR6	1.67e-43
40	390	14.0	345	11	O70129 ANAPHYLATOXIN C5A RECE	7.70e-43
41	375	13.5	375	4	O43494 G PROTEIN-COUPLED RECE	2.34e-40
42	359	12.9	334	11	O61125 BRADYKININ RECEPTOR, B	1.01e-37
43	351	12.6	168	11	O35715 SOMATOSTATIN RECEPTOR	2.06e-36
44	336	12.1	361	6	O46685 ORPHAN G PROTEIN-COUPLE	5.73e-34
45	329	11.8	360	11	O70171 L-CCR	7.83e-33

ALIGNMENTS

RESULT 1  
ID 054689 PRELIMINARY; PRT; 367 AA.  
AC 054689;  
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE G PROTEIN-COUPLED RECEPTOR KV411.  
OS MUS MUSCULUS (MOUSE).  
OC EKARFOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
RN EUTHERIA; RODENTIA.  
RC [1]  
RP SEQUENCE FROM N.A.  
RA YANAGIHARA S., KOMURA E., YAMAGUCHI Y.;  
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AB009369; D1024681;  
SQ SEQUENCE 367 AA; 42102 MW; 64D91F19 CRC32;  
Query Match 73.8%; Score 2051; DB 11; Length 367;  
Best Local Similarity 78.8%; Pred. No. 0.00e+00;  
Matches 256; Conservative 43; Mismatches 25; Indels 1; Gaps 1;

Db	43	IAYSLICVGLGILNIMVMTFAFYKARSMTDVLNNAITDILFVLTLFPFWAVTHATNT	102
Qy	42	IAYSLICVGLGILNIMVMTFAFYKARSMTDVLNNAITDILFVLTLFPFWAVSHATGA	101
Db	103	WVFSALCKLMKGYANFNCGMLLACISMDRYIAIQAATKSPRVSRTTHSHKVICVA	162
Qy	102	WVFSNATCKLKGIAYNFNCGMLLTCISMDRYIAIQAATKSPRLSRITLPRSKIIICLV	161
Db	163	WVFSIISSFTTFNKKYELQDRDCEPRYSRVSSEPTWKLGLGMLFGFTPLIFM	222
Qy	162	WVGLSVIISSFTTFNKKYELQDRDCEPRYSRVSSEPTWKLGLGMLFGFTPLIFM	221
Db	223	VFCVLIKTIVQAQNSKRHRRAIRVIAVLVAVLACQIPHNVLIVTAVNTGKVRSCST	282
Qy	222	IFCTVFIKTLVQAQNSKRHRRAIRVIAVLVAVLACQIPHNVLIVTAVNTGKVRSCQS	281
Db	283	EKVLATRNVAELAFHCCCLNPLYAFIGQFRNYFNKIMKDWCMRRKMKMFGFLCAR	342
Qy	282	EKLIGYKTKTEVLAFLHCCCLNPLYAFIGQFRNYFNKIMKDWCMRRKMKMFGFLCAR	341
Db	343	YISESYISROTSETVENDNASSFTM	367
Qy	342	RYSEN-ISROTSETADNDNASSFTM	365



```

RESULT 2
ID O42444 PRELIMINARY; PRT; 368 AA.
AC O42444;
DT 01-JAN-1998 (TREMREL. 05, CREATED)
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DE INTERLEUKIN-8-LIKE RECEPTOR.
OS CORONCHUS MYKISS (RAINBOW TROUT) (SALMO GAIREDNERI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA.
OC OSTEICHTHYES; ACTINOPTERYGII; SALMONIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RA ZOU J., DANIELS G.D., CUNNINGHAM C., SECOMBES C.J.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AJ003159; E1193478; -
DR PFAM; PF00001; 7tm_1.
SQ SEQUENCE 368 AA; 41523 MW; 56D3903B CRC32;

Query Match 30.0%; Score 833; DB 13; Length 368;
Best Local Similarity 41.1%; Pred. No. 1.60e-119;
Matches 116; Conservative 76; Mismatches 86; Indels 4; Gaps 4;

Db 54 YWSVILGGNLVWVWYILHFRORLKTMTDIYLLNLAVALDFLGLTLPWAV-EANQGW 112
| : ||||| : | : | : ||||| : ||||| : ||||| : | : |
QY 44 YSLICVLGLGNILV-ITFAFYKARSMTDYLNNMAIADILFVLTLPFWAVSHATGAW 102
| : ||||| : | : | : ||||| : ||||| : ||||| : | : |

Db 113 SFGLGCKVTSAFYKINFFSMLLTICISVDRIYVIVQTTMAQNSKROSLSCSKFVCACV 172
| : || : ||||| : ||||| : ||||| : ||||| : ||||| : | : |
QY 103 VFSNATCKLLKGIYAINFCNGMLLTICISMDRYAIVQATSKFSRLSRTPRSKIICLVV 162
| : ||||| : | : | : ||||| : ||||| : ||||| : | : |

Db 173 WLIAVLALPEFMFANKVELDQGYCTMYWS-NONNRTKIVVLGLQICMGFCPLPLVMV 231
| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : | : |
QY 163 WGLSVIISSTFVNQYNTQSDVCEPKYQTVSEPIRWKLLMLGLLELFGFFPLMFEM 222
| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : | : |

Db 232 FCYAGIIRTLLKTSFOKKALRYLVVAVFVLSQLPYNSVLMVEATQANSTOTDCA 291
| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : | : |
QY 223 FCYTFIVKTLVQAQNSKRHRKAIIRVIAVVLVFLACQIPHNMYLLTAANLGMNRS-CQS 281
| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : | : |

Db 292 AKRNVVSQVILKSLAYTHACLNPFLYFVGVRFRDILKLLR 333
| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : | : |
QY 282 EKLIGYTKTVTEVLAFLHCLNPLVYAFIGQKFRNYFLKILK 323
| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : | : |

RESULT 3
ID O55169 PRELIMINARY; PRT; 359 AA.
AC O55169;
DT 01-JUN-1998 (TREMREL. 06, CREATED)
DT 01-JUN-1998 (TREMREL. 06, LAST SEQUENCE UPDATE)
DE RECEPTOR PROTEIN CKR3.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA.
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA HARRINGTON P.M., NEWTON D.J., COLEMAN J.W., FLANAGAN B.F.;
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Y13400; E1247073; -
DR PFAM; PF00001; 7tm_1.
SQ SEQUENCE 359 AA; 41609 MW; 874DF714 CRC32;

Query Match 29.7%; Score 824; DB 11; Length 359;
Best Local Similarity 37.7%; Pred. No. 6.32e-118;
Matches 106; Conservative 91; Mismatches 76; Indels 8; Gaps 7;

Db 45 YSLVIFVGLGNMMVILIKRYKLQIMNTIYLLNLAISDLFLTFVPFW-IHYVLNWEW 103
| : ||||| : || : | : ||||| : ||||| : ||||| : | : |
QY 44 YSLICVLGLGNILVITFAFYKARSMTDYLNNMAIADILFVLTLPFWAVSHAT-GAW 102
| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : | : |

Db 104 GFCHCMCKMSGLYLYALYSEIFILLITIDRYLAIVAV--LAIRATVTFATTSIT 161
| : ||||| : || : | : ||||| : ||||| : ||||| : | : |

```



AC	062745;		
DT	01-AUG-1998 (TREMBREL. 07, CREATED)		
DT	01-AUG-1998 (TREMBREL. 07, LAST SEQUENCE UPDATE)		
DT	01-AUG-1998 (TREMBREL. 07, LAST ANNOTATION UPDATE)		
DE	CHEMOKINE RECEPTOR CCR5.		
OS	CERCOCEBUS TORQUATUS ATYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;		
OC	EUTHERIA; PRIMATES.		
[1]			
RN	SEQUENCE FROM N.A.		
RA	CHEN Z., GETTIE A., HO D.D., MARX P.A.;		
RL	VIROLOGY 0:0-0(1998).		
DR	EMBL; AF051905; G3135302; -.		
QY	SEQUENCE 352 AA; 40489 MW; PB9CE731 CRC32;		
Query Match 28.8%; Score 799; DB 6; Length 352;			
Best Local Similarity 38.9%; Pred. No. 1.70e-113;			
Matches 109; Conservative 87; Mismatches 73; Indels 11; Gaps			
Db	37 YSLVFIQFVGNILVLLINCKRLKSMTDIYLLNLAIISDLFLTLTFVFWA-HYAAQWD 95		
QY	44 YSLICVLGLGNILVITFAFYKARSMTDVLLNLMAADILFVLTLPFWAVSHATGAW 103		
Db	96 FGNMTCOLLTVGYTIGFPGSIFFIILLTDYLAIVHAV--FALKARVTFTGLVTSVITW 153		
QY	104 FSNATCKLLKGIYAINFCNGMLLTCTISMDRYIAIVQATKFSRLSRPLPSKIIICLVW 163		
Db	154 VVAVFASLPGIIFTRS-QREGLHYTCSPPHYSQYQF-WKNFOTIKIVIL-GLVLPLVM 210		
QY	164 GLSVIISSTSEVFNQYNTQSGD-VCEPKYQTVSEPIRWK-LLMLGLELLFGFPIPLMF 221		
Db	211 VICVSGILKILLRCNKKRRAVRLFTIMIVTFLFWAPNIVILLNTFOEFGFN-NC 269		
QY	222 IFCYTFIVKTLVQANSK-RHKAIRVIAVVLVFLACQIPHNMVLLVTA-ANLGNMNSC 279		
Db	270 SSSNRLQDQMOVTELTGMTCCINPIIVAFVGEKFRNVL 309		
QY	280 QSEKLGITKTVTEVLAFLHCLLNPLVYAFIQKFRNYFL 319		
RESULT	8 PRELIMINARY; PRT; 352 AA.		
ID	062745		
AC	062745;		
DT	01-AUG-1998 (TREMBREL. 07, CREATED)		
DT	01-AUG-1998 (TREMBREL. 07, LAST SEQUENCE UPDATE)		
DT	01-AUG-1998 (TREMBREL. 07, LAST ANNOTATION UPDATE)		
DE	CHEMOKINE RECEPTOR CCR5.		
OS	CERCOCEBUS TORQUATUS ATYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;		
OC	EUTHERIA; PRIMATES.		
[1]			
RN	SEQUENCE FROM N.A.		
RA	CHEN Z., GETTIE A., HO D.D., MARX P.A.;		
RL	VIROLOGY 0:0-0(1998).		
DR	EMBL; AF051904; G3135300; -.		
QY	SEQUENCE 352 AA; 40460 MW; E6A5AA07 CRC32;		
Query Match 28.7%; Score 798; DB 6; Length 352;			
Best Local Similarity 37.9%; Pred. No. 2.55e-113;			
Matches 113; Conservative 92; Mismatches 80; Indels 13; Gaps			
Db	19 PCQKINGQIAARLLPPLYSLVFIQFVGNLVLLINCKRLKSMTDIYLLNLAIISDLL 78		
QY	28 PCRR-SGSS-PGIYLRATYSLICVLGLGNILVITAFYKARSMTDVLLNLMAADIL 85		
Db	79 FLTLVPFWA-HYAAAQWDFGNMTCOLLGLTYFIFGFSGIFILLTDYLAIVHAV--F 135		
QY	86 FVLTLPEWAVSHATGAWVESNATCKLLKGIYAINFCNGMLLTCTISMDRYIAIVQATKSF 145		
Db	136 ALKARTVFGVTSVITWVAVFASLPGIIFTRS-QREGLHYTCSPPHYSQYQF-WKNE 193		
QY	146 RLRSRTLPRLKIIICLVYMGSLVSISSSTFVFNQYNTQSGD-VCEPKYQTVSEPIRWK-L 203		



```

Db 194 QTLKIVIL-GLVPLLVWICYSGILTKLRCKNEKRHRVRLITIMIVYFLFWAPYN 252
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 204 LMLGLELLFGFFIPMFIMFICFTYFIVATLVQAQNSK-RHKRAIRVIAVLVFLACQIPH 262

Db 253 IVLLNLTFOFFGLN-NCSSNRLDQAMQVETELGHTHCINIIYAFVGEKFRNYLL 309
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 263 MVLVUTA-ANLGMKNRSCQSEKLGYTKTVEVLAFLHCCLNPLVYAFIGOKFRNYFL 319

RESULT      9
ID O62744      PRELIMINARY;      PRG;      352 AA.
AC O62744;
DT 01-AUG-1998 (TREMBUREL. 07, CREATED)
DT 01-AUG-1998 (TREMBUREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBUREL. 07, LAST ANNOTATION UPDATE)
DE CHEMOKIN RECEPTOR CCR5.
OS CERCOBESUS TORQUATUS ATYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN Z., GETTIE A., HO D.D., MARX P.A.;
RL VIROLOGY 0:0-0(1998)
DR EMBL; AF051903; G3135298; -.
SQ SEQUENCE 352 AA; 40503 MW; C922372D CRC32;

Query Match      28.6%; Score 796; DB 6; Length 352;
Best Local Similarity 38.9%; Pred. No. 5.76e+113;
Matches 109; Conservative 87; Mismatches 73; Indels 11; Gaps 10;

Db 37 YSLVFIYFGVGNILVWLILINCLRLKASMTDYLLNLNAISDLLELLVTPFWA-HYAAAQWD 95
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 44 YSLICVLGLGNILVWITPFAFYKKARSMTDYLLNMAIDILVLTLPFWAVSHATGANV 103

Db 96 FGNMTCOLLTGLYFIFGFSGIFPIILLTIDRYAIAVAV--FALKARTVTFGLVTSVITW 153
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 104 FSNATCKLLKGIYAINFNCGMLLTCTISMDRYAIVQATKSFRLSRITLPRSKIIICLVV 163

Db 154 VVAVEASLPGIITRS-QRSLGHTVCSPPHYSOYF-WKNEQTAKIVIL-GLVPLLVW 210
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 164 GLSVITSSSTVFVNQKYNTOGSD-VCEPKYQTVSEPIRWK-LMLGLELLFGFFIPMF 221

Db 211 VICYSGLKTLKLRCKNEKRHRVRLITIMIVYFLFWAPYNLTGTFQFFGLN-NC 269
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 222 IFCVTFIVKTLVQAQNSK-RHKRAIRVIAVLVFLACQIPHNNVLLVTA-ANLGMKNRSC 279

Db 270 SSNRLDQAMQVETELGHTHCINIIYAFVGEKFRNYLL 309
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 280 QSEKLGYTKTVEVLAFLHCCLNPLVYAFIGOKFRNYFL 319

```

<hr/>						
RESULT	10					
ID	Q18770	PRELIMINARY;	PRT;	352 AA.		
AC	O18770;					
DT	01-JAN-1998	(TREMBREL. 05, CREATED)				
DT	01-JAN-1998	(TREMBREL. 05, LAST SEQUENCE UPDATE)				
DE	01-JUN-1998	(TREMBREL. 06, LAST ANNOTATION UPDATE)				
DE	CCR5 RECEPTOR (FRAGMENT).					
GN	CCR5..					
OS	FAN TROGLODYTES (CHIMPANZEE).					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;					
OC	EUTHERIA; PRIMATES.					
RN	{1}					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=WACCR5-140A.					
RA	ZHANG L., CARRUTHERS C.D., HE T., HUANG Y., CAO Y., WANG B.,					
RA	HO D.D.;					
RL	AIDS RES. HUM. RETROVIRUSES 0:0-0(1997).					
DR	EMBL; AF011538; G2305194; -.					
DR	PFAM; PF00001; 7tm_1.					
FT	NON_TER 352 352					
SQ	SEQUENCE 352 AA; 40523 MW; FLC10E99 CRC32:					



```

RESULT 12
ID O55193 PRELIMINARY; PRT; 373 AA.
AC O55193;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE CHEMOKINE RECEPTOR CCR2.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE DAWLEY;
RA JIANG Y., SALAFRANCA M.N., ADHIKARI S., XIA Y., FENG L., SONNTAG M.K.,
RA DEIERER C.M., PENNELL N.A., STREIT W.J., HARRISON J.K.;
RL J. NEUROIMMUNOL. 0:0-0(1998).
DR EMBL; U77349; G2896818; -.
SQ SEQUENCE 373 AA; 42763 MW; 14578A08 CRC32;

Query Match 27.8%; Score 772; DB 11; Length 373;
Best Local Similarity 37.7%; Pred. No. 1.0le-108;
Matches 113; Conservative 90; Mismatches 81; Indels 16; Gaps 14;

Db 44 PCHKTSVKQIGAWLPPPLYSLSVIFGFGVGNMLVLIILICKLKLSMTDYLFLNLAISDLL 103
QY 28 PCRRSGSP-G-YLYRIAYSLICVLGLGNILVITFAFYKKARSMVDYLLNMAIADIL 85

Db 104 FLLTLPFWA-HYAANEWFEGNICKLFTGLYHIGYFGGIFILLTIDRYLAIVHAV--F 160
QY 86 FVLTPFWAVSHATGAWFVSNAATCKLLKGIYAINFCNGLMLLTCSMDRYIAIVQATKSF 145

Db 161 ALKARTVFGVITSVTVVAVFASLPGIIFTKSEDDQHTGCP-Y--FP-TI-WKNFQ 215
QY 146 RLRSRTLPKRSKILICLVWGLSVIISSTVFNQKYNQTSQDVCEPKYQTVSEPIRWKLM 205

Db 216 TIMRNLSILPLVWVICYSGLHFTPCRNKKRRAVRLIFAMIVYFLEWTPVNIY 275
QY 206 LGLELFGFFLPMFIMFYCTIVTKLVQAQNSK-RHKAIRVIAVVLVFLACQIPHNV 264

Db 276 LFLATTFQBLG-MS-NCVDMHLDQAMQVETLGMTHCCVNPYIYAFVGEKFRY-LSIF 332
QY 265 LLVTA-AN-LGKMNRSQCEKLGITKTVTEVLAFHCLNPLVYAFIGQKFRNYFLKIL 322

RESULT 13
ID O18772 PRELIMINARY; PRT; 352 AA.
AC O18772;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE CCR5 RECEPTOR (FRAGMENT).
GN CCR5.
OS PAN TROGLODYTES (CHIMPANZEE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHCRR5-142A;
RA ZHANG L., CARRUTHERS C.D., HE T., HUANG Y., CAO Y., WANG G., HAHN B.,
RA HO D.D.;
RL AIDS RES. HUM. RETROVIRUSES 0:0-0(1997).
DR EMBL; AF011541; G2305200; -.
DR PFAM; PF00001; 7tm1.
FT NON_TER 352
SQ SEQUENCE 352 AA; 40598 MW; A9BF8EDF CRC32;

Query Match 27.7%; Score 769; DB 6; Length 352;
Best Local Similarity 36.8%; Pred. No. 3.42e-108;
Matches 103; Conservative 92; Mismatches 74; Indels 11; Gaps 10;

Db 37 YSLVIFGFGVGNMLVILILINKRKLSMTDYLFLNLAISDLTFLVFPFWA-HYAAAQWD 95
QY 44 YSLICVLGLGNILVITFAFYKKARSMVDYLLNMAIADILFVLTPFWAVSHATGAW 103

Db 96 FGNTMCOLLGLYFIFGFGGIFILLTIDRYLAIVHAV--FALKARTVTFGVTSVITW 153
QY 104 FSNATCKLLKGIYAINFCNGLMLLTCSMDRYIAIVQATKSFRLRSRTLPKRSKILCVW 163

Db 154 VVAVFASLPGIIFTRS-QKEGLHYTCSHFPYSQVF-WKNFQTLKIVIL-GLVPLLVW 211
QY 164 GLSVIISSTVFNQKYNQTSQDVCEPKYQTVSEPIRWK-LMLGLELFGFFLPMFIM 222

Db 212 ICYSGILKTLILRCNKKRRAVRLIFTMIVYFLEWAPYINVLNLTNFOQFFGLN-NCS 270
QY 223 FCTYFIVTKLVQAQNSK-RHKAIRVIAVVLVFLACQIPHNVLLVTA-ANLGMNRSCQ 280

Db 271 SSNRLDQAMQVETLGMTHCCINPIIYAFVGEKFRNYLL 309
QY 281 SEKLIGYTKTVTEVLAFHCLNPLVYAFIGQKFRNYFL 319

RESULT 15
ID O00537 PRELIMINARY; PRT; 384 AA.
AC O00537;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE CHEMOKINE RECEPTOR CCR-9.
OS HOMO SAPIENS (HUMAN).

```

```

Db 96 FGNTMCOLLGLYFIFGFGGIFILLTIDRYLAIVHAV--FALKARTVTFGVTSVITW 153
QY 104 FSNATCKLLKGIYAINFCNGLMLLTCSMDRYIAIVQATKSFRLRSRTLPKRSKILCVW 163

Db 154 VVAVFASLPGIIFTRS-QKEGLHYTCSHFPYSQVF-WKNFQTLKIVIL-GLVPLLVW 210
QY 164 GLSVIISSTVFNQKYNQTSQDVCEPKYQTVSEPIRWK-LMLGLELFGFFLPMFIM 221

Db 211 VICYSGILKTLILRCNKKRRAVRLIFTMIVYFLEWAPYIDIVLLNLTNFOQFFGLN-NC 269
QY 222 FCTYFIVTKLVQAQNSK-RHKAIRVIAVVLVFLACQIPHNVLLVTA-ANLGMNRSC 279

Db 270 SSNRLDQAMQVETLGMTHCCINPIIYAFVGEKFRNYLL 309
QY 280 SEKLIGYTKTVTEVLAFHCLNPLVYAFIGQKFRNYFL 319

```

```

RESULT 14
ID O18771 PRELIMINARY; PRT; 352 AA.
AC O18771;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE CCR5 RECEPTOR (FRAGMENT).
GN CCR5.
OS PAN TROGLODYTES (CHIMPANZEE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHCRR5-141A;
RA ZHANG L., CARRUTHERS C.D., HE T., HUANG Y., CAO Y., WANG G., HAHN B.,
RA HO D.D.;
RL AIDS RES. HUM. RETROVIRUSES 0:0-0(1997).
DR EMBL; AF011539; G2305196; -.
DR PFAM; PF00001; 7tm1.
FT NON_TER 352
SQ SEQUENCE 352 AA; 40466 MW; D52C67E1 CRC32;

Query Match 27.7%; Score 769; DB 6; Length 352;
Best Local Similarity 36.9%; Pred. No. 3.42e-108;
Matches 103; Conservative 92; Mismatches 75; Indels 9; Gaps 8;

```

```

Db 37 YSLVIFGFGVGNMLVILILINKRKLSMTDYLFLNLAISDLTFLVFPFWA-HYAAAQWD 95
QY 44 YSLICVLGLGNILVITFAFYKKARSMVDYLLNMAIADILFVLTPFWAVSHATGAW 103

Db 96 FGNTMCOLLGLYFIFGFGGIFILLTIDRYLAIVHAV--FALKARTVTFGVTSVITW 153
QY 104 FSNATCKLLKGIYAINFCNGLMLLTCSMDRYIAIVQATKSFRLRSRTLPKRSKILCVW 163

Db 154 VVAVFASLPGIIFTRS-QKEGLHYTCSHFPYSQVF-WKNFQTLKIVIL-GLVPLLVW 211
QY 164 GLSVIISSTVFNQKYNQTSQDVCEPKYQTVSEPIRWK-LMLGLELFGFFLPMFIM 222

Db 212 ICYSGILKTLILRCNKKRRAVRLIFTMIVYFLEWAPYINVLNLTNFOQFFGLN-NCS 270
QY 223 FCTYFIVTKLVQAQNSK-RHKAIRVIAVVLVFLACQIPHNVLLVTA-ANLGMNRSCQ 280

Db 271 SSNRLDQAMQVETLGMTHCCINPIIYAFVGEKFRNYLL 309
QY 281 SEKLIGYTKTVTEVLAFHCLNPLVYAFIGQKFRNYFL 319

RESULT 15
ID O00537 PRELIMINARY; PRT; 384 AA.
AC O00537;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE CHEMOKINE RECEPTOR CCR-9.
OS HOMO SAPIENS (HUMAN).

```







(TM)

```
Run on:      Fri Nov 13 12:08:17 1998;  MasPar time 4.12 Seconds
          627.819 Million cell updates/sec
Tabular output not generated
```

Result No.	Query Match	Score	Query			DB	ID	Description	Pred. No.
			Length	DB	ID				
1	984	35.4	350	2	PCT-US95-0	Sequence 7	Applicatio	8.26e-79	
2	984	35.4	350	1	US-08-701-	Sequence 2	Applicatio	8.26e-79	
3	984	35.4	350	1	US-08-410-	Sequence 1	Applicatio	8.26e-79	
4	984	35.4	350	1	US-08-076-	Sequence 2	Applicatio	8.26e-79	
5	984	35.4	350	1	US-08-203-	Sequence 1	Applicatio	8.26e-79	
6	984	35.4	350	1	US-08-410-	Sequence 1	Applicatio	8.26e-79	
7	984	35.4	350	1	US-08-450-	Sequence 7	Applicatio	8.26e-79	
8	961	34.6	358	2	PCT-US93-1	Sequence 19	Applicati	1.15e-76	
9	961	34.6	358	1	US-08-153-	Sequence 19	Applicati	1.15e-76	
10	961	34.6	378	2	PCT-US93-1	Sequence 15	Applicati	1.15e-76	
11	961	34.6	378	1	US-08-153-	Sequence 15	Applicati	1.15e-76	
12	961	34.6	410	1	US-08-153-	Sequence 7	Applicatio	1.15e-76	
13	961	34.6	410	2	PCT-US93-1	Sequence 7	Applicatio	1.15e-76	
14	951	34.2	355	2	PCT-US95-0	Sequence 8	Applicatio	9.77e-76	
15	951	34.2	355	1	US-07-759-	Sequence 1	Applicatio	9.77e-76	
16	951	34.2	355	1	US-08-450-	Sequence 8	Applicatio	9.77e-76	
17	951	34.2	360	1	US-08-202-	Sequence 7	Applicatio	9.77e-76	
18	948	34.1	359	1	US-08-153-	Sequence 24	Applicati	1.86e-75	
19	948	34.1	359	2	PCT-US93-1	Sequence 24	Applicati	1.86e-75	
20	928	33.4	378	1	US-08-383-	Sequence 2	Applicatio	1.35e-73	
21	928	33.4	378	1	US-08-383-	Sequence 2	Applicatio	1.35e-73	
22	928	33.4	378	2	PCT-US93-0	Sequence 2	Applicatio	1.35e-73	
23	906	32.6	312	1	US-08-118-	Sequence 38	Applicati	1.50e-71	











[illegible]

RESULT	6
ID	US-08-410-453A-1 STANDARD; PRT; 350 AA.
XX	XXXXXX
AC	
DT	
DE	Sequence 1, Application US/08410453A
CC	Sequence 1, Application US/08410453A
CC	Patent No. 5767063
CC	GENERAL INFORMATION:
CC	APPLICANT: Lee, James,
CC	APPLICANT: Holmes, William E.,
CC	APPLICANT: Woods, William I.,
CC	TITLE OF INVENTION: Human PF4A Receptors and Their Use
CC	NUMBER OF SEQUENCES: 2
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Genentech, Inc.
CC	STREET: 1 DNA Way
CC	CITY: South San Francisco
CC	STATE: California
CC	COUNTRY: USA
CC	ZIP: 94080
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: Winpatin (Genentech)
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/410,453A
CC	FILING DATE: 24-Mar-1995
CC	CLASSIFICATION: 514
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: 08/234494
CC	FILING DATE: 28-APR-1994
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: 07/672711
CC	FILING DATE: 29-MAR-1991
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Love, Richard B.
CC	REGISTRATION NUMBER: 34,659
CC	REFERENCE/DOCKET NUMBER: P0706C1D1
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: 650/225-5530
CC	TELEFAX: 650/953-9881
CC	INFORMATION FOR SEQ ID NO: 1:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 350 amino acids
CC	TYPE: Amino Acid
CC	TOPOLOGY: Linear
SQ	SEQUENCE 350 AA; 39805 MW; 660082 CN;
	Query Match 35.4%; Score 984; DB 1; Length 350;
	Best Local Similarity 44.4%; Pred. No. 8.26e-79;
	Matches 134; Conservative 76; Mismatches 80; Indels 1
Dd	27 Y-SPCMLETEFLNKYVWIIAYALVFLLSLGNSLYMWILYSRVGRSVTDVYL
Qy	25 YCAPCRSSSGPG-YLYIRIAYSLICVLGLGNILVIITFAFYKARSMTDVLL
Dd	86 LLFALTLPWAASKVNG-WIFGTFLCKVWSLLKEVNFYSGILLACISVDRLYA
Qy	84 ILFVLTLPFWAVSHATGAWFVSNAACKLKLGIYAINFCMMLLTCTISMDRYIA
Dd	145 TLTK-RHL--VKFYCLGCWGLSNLSLPPFLFRQAYPNNSSPYC---EYELGG
Qy	144 SFLRSRTLPRSKIIICLVWGSLVSISSTSFEVNKYNTQGS-D-VCEPKYQTVS
Dd	199 RMVLRLPHITGFIVPLFVMLFCYGFTLRTLFKAHMGOKHMRVIFAWVLFL
Qy	202 KLLMLGELLGEFFPLMFIMFFIKVTLLVQAQNSRRHRAIRVIIIAVLVLF



Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGFLHSCNLPIIYAFIQGNFRHGFLK 318  
QY 262 NMVLLVTAAANLGRMNR-SCQSEKLIQYTKVTVEVLAFLHCLNPLVYAFIQGKFRNYFLK 320  
Db 319 IL 320  
QY 321 IL 322  
RESULT 7  
ID US-08-450-393A-7 STANDARD; PRT; 350 AA.  
XX  
AC xxxxxx  
DT  
XX  
XX  
DE Sequence 7, Application US/08450393A  
XX  
XX Patent No. 5707815  
CC Patent US/08450393A  
CC  
CC GENERAL INFORMATION:  
CC APPLICANT: Charo, Israel  
CC APPLICANT: Coughlin, Shaun  
CC TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT  
CC TITLE OF INVENTION: PROTEIN RECEPTORS  
CC  
CC NUMBER OF SEQUENCES: 14  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
CC STREET: 5 Palo Alto Square  
CC CITY: Palo Alto  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94306-2155  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/450,393A  
CC FILING DATE: May 25, 1995  
CC CLASSIFICATION: 424  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Cseri, Luann  
CC REGISTRATION NUMBER: 31,822  
CC REFERENCE/DOCKET NUMBER: UCAL-237/02US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 415-843-5165  
CC TELEFAX: 415-8857-0663  
CC TELEX: 380816COOLEYPA  
CC INFORMATION FOR SEQ ID NO: 7:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 350 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC HYPOTHETICAL: NO  
CC ANTI-SENSE: NO  
SQ SEQUENCE 350 AA; 39805 MW; 660082 CN;  
Query Match 35.4%; Score 984; DB 1; Length 350;  
Best Local Similarity 44.4%; Pred. No. 8.26e-79;  
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;  
Db 27 Y-SPCMETETLNKVVIIAYALVFLSLGNSLMLVILSRVGRSVTDVYLNLAAD 85  
QY 25 YCAPCRSSGSPG-YLYRIAYSLICVLGLNLVITFAFYKRSNTDYLNLMAIAD 83  
Db 86 LLFALPLTWAASKVNG-WLFGFLCKVWSLLKKEVNFYSGILLIACISVDYRLAIVHATR 144  
QY 84 ILFVLTPWAVSHAGAWFNSNATCKLLKGIYAINFCGMLLTCISMDRIYIAVQATK 143  
Db 145 TLTKQ-RHL--VRFVCLGCGWGLSMNLSLPFFLFRQYHPNNSPVC---YEVLGNDTAKW 198

QY 144 SFPLRSRTLPKRIICLVVGLSVLISSTFVFNQYNTGSD-VCEPKYQTVS-EPINW 201  
Db 199 RMVLRILPHTFGTIVPLFVMLFCYGTFLTLFKAHMGQKHRAMRVIFAVVLIFLCLWLPY 258  
QY 202 KLMLGLELLFGFFIPLMFIMFYTFIVKTLVQAQSKRHKAIKRVIIAVVFLACQIPH 261  
Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGFLHSCNLPIIYAFIQGNFRHGFLK 318  
QY 262 NMVLLVTAAANLGRMNR-SCQSEKLIQYTKVTVEVLAFLHCLNPLVYAFIQGKFRNYFLK 320  
Db 319 IL 320  
QY 321 IL 322  
RESULT 8  
ID PCT-US93-11153-19 STANDARD; PRT; 358 AA.  
XX  
AC xxxxxx  
DT  
XX  
XX  
DE Sequence 19, Application PC/TUS9311153  
XX  
XX Sequence 19, Application PC/TUS9311153  
CC GENERAL INFORMATION:  
CC APPLICANT: Godiska, Ronald  
CC APPLICANT: Gray, Patrick W.  
CC APPLICANT: Schweikart, Vicki L.  
CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors  
CC NUMBER OF SEQUENCES: 84  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
CC ADDRESSEE: Bicknell  
CC STREET: 6300 Sears Tower, 233 South Wacker Drive  
CC CITY: Chicago  
CC STATE: Illinois  
CC COUNTRY: USA  
CC ZIP: 60606  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/11153  
CC FILING DATE:  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 07/977,452  
CC FILING DATE: 17-NOV-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Noland, Greta E.  
CC REGISTRATION NUMBER: 35,302  
CC REFERENCE/DOCKET NUMBER: 31794  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (312) 474-6300  
CC TELEFAX: (312) 474-0448  
CC TELEX: 25-3856  
CC INFORMATION FOR SEQ ID NO: 19:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 358 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 358 AA; 40690 MW; 701130 CN;  
Query Match 34.6%; Score 961; DB 2; Length 358;  
Best Local Similarity 45.7%; Pred. No. 1.15e-76;  
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;  
Db 40 FL-PIWISICFVLGLNGLVVLTYTFYFRKRLTKMTDTLLNLAVADILFLLLPFWAYS- 97







```
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 378 AA; 42873 MW; 782819 CN;
Query Match 34.6%; Score 961; DB 2; Length 378;
Best Local Similarity 45.7%; Pred. No. 1.15e-76;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;
Db 60 FL-PIMYSIIICFVGLGNGLVLTYYIFKRLKMTDTYLLNLAVALDILFLLTLPFWAYS- 117
QY 38 YLYRIAYSLICVLGNGLVLTYYIFKRLKMTDTYLLNLAVALDILFLLTLPFWAVSH 97
Db 118 AAKSWFVGHFCKLIFAIYKMSFSGMLLLCISIDRYVAIVQAVSAHRHRAVLLISKL 177
QY 98 ATGAWFVSNAATCKLKGIIYAINFCGMLLLTCISMDRYIAIVQAKSFRLSRTLPRSKI 157
Db 178 SCVGIWILATVSLPELSD-LQRSSEQ-AMRCSLITEHVE-AFITIQAQMVIGFLV 234
QY 158 ICLVWGLSVIISSTFVFNQKYNTOGSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216
Db 235 PLLAMSCYLVIRTLQARNFERNKAIVIAVVVFIQFQLPYNGVVLQAOTVANFNIT 294
QY 217 PLMFIMFYCTFIVKTLVQAQSKRKAIRVIAVVLVFLACQIPHNMVLLV-TAANLGM 275
Db 295 SSTCELSKQINAIYDVTYSLACVRCVNPFLYAFIGVFRNDLFLKFLKDLGCL 347
QY 276 NRSCKSEKLGITKTVTEVLAFLHCCLPVLYAFIGQKFRNFYFLKILKDLWCVC 328
RESULT 11
ID US-08-153-848-15 STANDARD; PRT; 378 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 15, Application US/08153848
XX Sequence 15, Application US/08153848
CC Patent No. 5759804
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald
CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.
CC TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/153,848
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5759804and, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31794
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 474-6300
CC TELEFAX: (312) 474-0448
CC
```

```
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 15:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 378 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 378 AA; 42873 MW; 782819 CN;
Query Match 34.6%; Score 961; DB 1; Length 378;
Best Local Similarity 45.7%; Pred. No. 1.15e-76;
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;
Db 60 FL-PIMYSIIICFVGLGNGLVLTYYIFKRLKMTDTYLLNLAVALDILFLLTLPFWAYS- 117
QY 38 YLYRIAYSLICVLGNGLVLTYYIFKRLKMTDTYLLNLAVALDILFLLTLPFWAVSH 97
Db 118 AAKSWFVGHFCKLIFAIYKMSFSGMLLLCISIDRYVAIVQAVSAHRHRAVLLISKL 177
QY 98 ATGAWFVSNAATCKLKGIIYAINFCGMLLLTCISMDRYIAIVQAKSFRLSRTLPRSKI 157
Db 178 SCVGIWILATVSLPELSD-LQRSSEQ-AMRCSLITEHVE-AFITIQAQMVIGFLV 234
QY 158 ICLVWGLSVIISSTFVFNQKYNTOGSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216
Db 235 PLLAMSCYLVIRTLQARNFERNKAIVIAVVVFIQFQLPYNGVVLQAOTVANFNIT 294
QY 217 PLMFIMFYCTFIVKTLVQAQSKRKAIRVIAVVLVFLACQIPHNMVLLV-TAANLGM 275
Db 295 SSTCELSKQINAIYDVTYSLACVRCVNPFLYAFIGVFRNDLFLKFLKDLGCL 347
QY 276 NRSCKSEKLGITKTVTEVLAFLHCCLPVLYAFIGQKFRNFYFLKILKDLWCVC 328
RESULT 12
ID US-08-153-848-7 STANDARD; PRT; 410 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 7, Application US/08153848
XX Sequence 7, Application US/08153848
CC Patent No. 5759804
CC GENERAL INFORMATION:
CC APPLICANT: Godiska, Ronald
CC APPLICANT: Gray, Patrick W.
CC APPLICANT: Schweikart, Vicki L.
CC TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
CC NUMBER OF SEQUENCES: 64
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Bicknell
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/153,848
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION NUMBER: US 07/977,452
CC FILING DATE: 17-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5759804and, Greta E.
```



CC REGISTRATION NUMBER: 35,302  
CC REFERENCE/DOCKET NUMBER: 31794  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (312) 474-6300  
CC TELEFAX: (312) 474-0448  
CC TELEX: 25-3856  
CC INFORMATION FOR SEQ ID NO: 7:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 410 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 410 AA; 46569 MW; 921741 CN;

Query Match 34.6%; Score 961; DB 1; Length 410;  
Best Local Similarity 45.7%; Pred. No. 1.15e-76;  
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;

Db 92 EL-PIMYSIIICFVGLLGNGLVLTYYFKRLKMTDTYLLNLAVADILFLTLPLFWAYS- 149  
QY 38 YLRYAISLCIVGLLGNLIVITFAFYKRSMTDVLNMAIDILFLTLPLFWAVSH 97  
Db 150 AAKSWVGFHCFKLIPIAIYKMSFFSGMLLLCISIDRYVAIVQAVSAHRHRAVLLISK 209  
QY 98 ATGANWFSNATCKLLKGIYAINFCNMLLTCTISMDRYIAIVQATKSFRLSRITLPRSKI 157  
Db 210 SCVGIWILATVISIPELLYSD-LQRSSEQ-AMRCSLITEHVE-AFTIQVQAVMIGFLV 266  
QY 158 ICLVWGLSVIISSTFVFNQKNTQSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216  
Db 267 PLLAMSFCLVLIIRTLLQARNFERNKAIRVIIVAVVVFVIFQLPYNGVLAQTVANENIT 326  
QY 217 PLMFIFCFYTFIVKTLVQAQNSKRHRKAIIRVIAVVLVFLACQIPHNNVLLV-TAANLGM 275  
Db 327 SSTCELSKOLNIAYDVITYSLACVRCVNPFLYAFIGVKFRNDLFKFLKDLGCL 379  
QY 276 NRSCQSEKLGITKTVTEVLAFLHCCNLPVLYAFIGKFRNFYFKILKDLWCV 328

RESULT 13  
ID PCT-US93-11153-7 STANDARD; PRT; 410 AA.  
XX XXXXX  
XX

Sequence 7, Application PC/TUS9311153

Sequence 7, Application PC/TUS9311153

GENERAL INFORMATION:  
CC APPLICANT: Godiska, Ronald  
CC APPLICANT: Gray, Patrick W.  
CC APPLICANT: Schweikart, Vicki L.  
CC TITLE OF INVENTION: Novel Seven Transmembrane Receptors  
CC NUMBER OF SEQUENCES: 64  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
CC ADDRESSEE: Bicknell  
CC STREET: 6300 Sears Tower, 233 South Wacker Drive  
CC CITY: Chicago  
CC STATE: Illinois  
CC COUNTRY: USA  
CC ZIP: 60606  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/11153  
CC FILING DATE:  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/977,452  
CC FILING DATE: 17-NOV-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Noland, Greta E.  
CC REGISTRATION NUMBER: 35,302  
CC REFERENCE/DOCKET NUMBER: 31794  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (312) 474-6300  
CC TELEFAX: (312) 474-0448  
CC TELEX: 25-3856  
CC INFORMATION FOR SEQ ID NO: 7:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 410 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 410 AA; 46569 MW; 921741 CN;

Query Match 34.6%; Score 961; DB 2; Length 410;  
Best Local Similarity 45.7%; Pred. No. 1.15e-76;  
Matches 134; Conservative 77; Mismatches 75; Indels 7; Gaps 7;

Db 92 EL-PIMYSIIICFVGLLGNGLVLTYYFKRLKMTDTYLLNLAVADILFLTLPLFWAYS- 149  
QY 38 YLRYAISLCIVGLLGNLIVITFAFYKRSMTDVLNMAIDILFLTLPLFWAVSH 97  
Db 150 AAKSWVGFHCFKLIPIAIYKMSFFSGMLLLCISIDRYVAIVQAVSAHRHRAVLLISK 209  
QY 98 ATGANWFSNATCKLLKGIYAINFCNMLLTCTISMDRYIAIVQATKSFRLSRITLPRSKI 157  
Db 210 SCVGIWILATVISIPELLYSD-LQRSSEQ-AMRCSLITEHVE-AFTIQVQAVMIGFLV 266  
QY 158 ICLVWGLSVIISSTFVFNQKNTQSDVCEPKYQTVSEPIRWKLLMLGL-ELLFGFFI 216  
Db 267 PLLAMSFCLVLIIRTLLQARNFERNKAIRVIIVAVVVFVIFQLPYNGVLAQTVANENIT 326  
QY 217 PLMFIFCFYTFIVKTLVQAQNSKRHRKAIIRVIAVVLVFLACQIPHNNVLLV-TAANLGM 275  
Db 327 SSTCELSKOLNIAYDVITYSLACVRCVNPFLYAFIGVKFRNDLFKFLKDLGCL 379  
QY 276 NRSCQSEKLGITKTVTEVLAFLHCCNLPVLYAFIGKFRNFYFKILKDLWCV 328

RESULT 14  
ID PCT-US95-00476-8 STANDARD; PRT; 355 AA.  
XX XXXXX  
XX

Sequence 8, Application PC/TUS9500476

Sequence 8, Application PC/TUS9500476

GENERAL INFORMATION:  
CC APPLICANT: The Regents of the University of California  
CC TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT  
CC NUMBER OF SEQUENCES: 14  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Robbins, Berliner & Carson  
CC STREET: 201 N. Figueroa Street, 5th Floor  
CC CITY: Los Angeles  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 90012-2628  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/00476  
CC FILING DATE:



CC CLASSIFICATION: 19910913  
CC ATTORNEY/AGENT INFORMATION: 435  
CC NAME: Berliner, Robert  
CC REGISTRATION NUMBER: 20,121  
CC REFERENCE/DOCKET NUMBER: 5555-291  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 310-977-1001  
CC TELEFAX: 310-977-1003  
CC TELEX:  
CC INFORMATION FOR SEQ ID NO: 8:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 355 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC HYPOTHETICAL: NO  
CC SEQUENCE 355 AA; 40122 MW; 676446 CN;

Query Match 34.2%; Score 951; DB 2; Length 355;  
Best Local Similarity 43.8%; Pred. No. 9,77e-76;  
Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;  
Db 44 YFVVIYALVFLLSLGNLSVLMVILYSRVGRSVTDVYLLNLALADLLFALTLPWAASK 103  
QY 38 YLYRIAYSLICVLGILGNILVITFAFYKARSMTDVYLLNMAIADILFVLTLPFWAVSH 97  
Db 104 VNG-WIFGFLCKVSVLLKEVNFYSIGILLACISVDRIYLAIVHATRTLTQK-RYL--VKF 159  
QY 98 ATGAWVFSNATCKLLKGIYAINFNGMLLTICISMDRYIAIVQATKSFRLSRTPRSKI 157  
Db 160 ICLTWGLSLLALPVLFRRTVSSNVSPAC---YEDMGNNTANWMLRLILPQSGFI 216  
QY 158 ICLVWGLSVIISSTFVFNQK-YNTQGSVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215  
Db 217 VPLLMLFCYGTLETLEKAMGQKRAMRVIFAVVLIFFLCWLPYNLVLLADTLMTQV 276  
QY 216 IPLMFIFCYTIVKTLVQAQNSKRKRAIRVIAVVLVFLACQIPHNMVLLVTAANLGM 275  
Db 277 IQETCERNHIDRALDATEILGILHSCNPLIYAFIQGKFRHGLKIL 324  
QY 276 NR-SCQSEKLGITKVTVEVLAFLHCCLNPNVLYAFIQGKFRNFKIL 322

RESULT 15  
ID US-07-759-568-1 STANDARD; PRT; 355 AA.  
XX  
AC xxxxxx  
XX

Sequence 1, Application US/07759568  
Sequence 1, Application US/07759568  
Patent No. 5374506  
GENERAL INFORMATION:  
APPLICANT: Murphy, Philip M.  
TITLE OF INVENTION: Cloning of cDNA Encoding a Functional  
TITLE OF INVENTION: Human Interleukin-8 Receptor  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cushman, Darby & Cushman  
STREET: 1615 L Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07759,568

CC FILING DATE: 19910913  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Scott, Watson T.  
CC REGISTRATION NUMBER: 26581  
CC REFERENCE/DOCKET NUMBER: WTS/5683/91535/WBH  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 202-861-3000  
CC TELEFAX: 202-822-0944  
CC TELEX: 6714627 cush  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 355 amino acids  
CC TYPE: AMINO ACID  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 355 AA; 40122 MW; 676446 CN;  
Query Match 34.2%; Score 951; DB 1; Length 355;  
Best Local Similarity 43.8%; Pred. No. 9,77e-76;  
Matches 126; Conservative 80; Mismatches 72; Indels 10; Gaps 7;  
Db 44 YFVVIYALVFLLSLGNLSVLMVILYSRVGRSVTDVYLLNLALADLLFALTLPWAASK 103  
QY 38 YLYRIAYSLICVLGILGNILVITFAFYKARSMTDVYLLNMAIADILFVLTLPFWAVSH 97  
Db 104 VNG-WIFGFLCKVSVLLKEVNFYSIGILLACISVDRIYLAIVHATRTLTQK-RYL--VKF 159  
QY 98 ATGAWVFSNATCKLLKGIYAINFNGMLLTICISMDRYIAIVQATKSFRLSRTPRSKI 157  
Db 160 ICLTWGLSLLALPVLFRRTVSSNVSPAC---YEDMGNNTANWMLRLILPQSGFI 216  
QY 158 ICLVWGLSVIISSTFVFNQK-YNTQGSVCEPKYQTVSE-PIRWKLLMLGLELFGFF 215  
Db 217 VPLLMLFCYGTLETLEKAMGQKRAMRVIFAVVLIFFLCWLPYNLVLLADTLMTQV 276  
QY 216 IPLMFIFCYTIVKTLVQAQNSKRKRAIRVIAVVLVFLACQIPHNMVLLVTAANLGM 275  
Db 277 IQETCERNHIDRALDATEILGILHSCNPLIYAFIQGKFRHGLKIL 324  
QY 276 NR-SCQSEKLGITKVTVEVLAFLHCCLNPNVLYAFIQGKFRNFKIL 322

Search completed: Fri Nov 13 12:08:42 1998  
Job time : 25 secs.



\*\*\*\*\*

W P S R E H (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Nov 13 12:09:01 1998; MasPar time 20.16 Seconds  
Tabular output not generated. 666.150 Million cell updates/sec

Title: >US-08-887-977-10  
Description: (1-365) from US08887977.pap  
Perfect Score: 2779  
Sequence: 1 MFSTPVKILQCSILHTQL.....NISQTSADNDNASSFTM 365

Scoring table: PAM 150  
Gap 11

Searched: 313962 seqs, 36788388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-pending  
1:P9 2:U60 3:U7 4:U80 5:U81 6:U82 7:U83 8:U84 9:U85  
10:U86 11:U87 12:U88 13:U89 14:U90 15:U91 16:NEWP  
17:NEWU6 18:NEWU7 19:NEWU8 20:NEWU9

Statistics: Mean 35.183; Variance 169.443; scale 0.208

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2779	100.0	365	10	US-08-675-Sequence 8, Applicatio	1.60e-231
2	2779	100.0	365	12	US-08-887-Sequence 10, Applicati	1.60e-231
3	2445	88.0	374	14	US-09-045-Sequence 48, Applicati	9.38e-202
4	984	35.4	350	12	US-08-802-Sequence 2, Applicatio	1.72e-72
5	984	35.4	350	15	US-09-104-Sequence 2, Applicatio	1.72e-72
6	984	35.4	350	8	US-08-450-Sequence 7, Applicatio	1.72e-72
7	984	35.4	350	12	US-08-801-Sequence 2, Applicatio	1.72e-72
8	984	35.4	350	6	US-08-284-Sequence 2, Applicatio	1.72e-72
9	984	35.4	350	1	PCT-US93-1Sequence 3, Applicatio	1.72e-72
10	984	35.4	350	12	US-08-805-Sequence 2, Applicatio	1.72e-72
11	984	35.4	350	8	US-08-410-Sequence 1, Applicatio	1.72e-72
12	984	35.4	350	7	US-08-308-Sequence 3, Applicatio	1.72e-72
13	984	35.4	350	8	US-08-441-Sequence 2, Applicatio	1.72e-72
14	984	35.4	350	12	US-08-801-Sequence 2, Applicatio	1.72e-72
15	984	35.4	350	12	US-08-802-Sequence 2, Applicatio	1.72e-72
16	984	35.4	350	8	US-08-446-Sequence 7, Applicatio	1.72e-72
17	984	35.4	350	15	US-09-104-Sequence 2, Applicatio	1.72e-72
18	984	35.4	350	5	US-08-182-Sequence 7, Applicatio	1.72e-72
19	984	35.4	1060	10	US-08-605-Sequence 21, Applicati	1.72e-72
20	961	34.6	358	6	US-08-245-Sequence 19, Applicati	1.75e-70
21	961	34.6	378	6	US-08-245-Sequence 15, Applicati	1.75e-70

22	961	34.6	410	6	US-08-245-Sequence 7, Applicatio	1.75e-70
23	961	34.6	410	3	US-07-977-Sequence 7, Applicatio	1.75e-70
24	959	34.5	378	6	US-08-245-Sequence 66, Applicati	2.62e-70
25	951	34.2	353	1	PCT-US96-0Sequence 9, Applicatio	1.31e-69
26	951	34.2	353	1	PCT-US96-0Sequence 8, Applicatio	1.31e-69
27	951	34.2	355	8	US-08-446-Sequence 8, Applicatio	1.31e-69
28	951	34.2	355	1	PCT-US93-1Sequence 4, Applicatio	1.31e-69
29	951	34.2	355	5	US-08-182-Sequence 8, Applicatio	1.31e-69
30	951	34.2	355	8	US-08-441-Sequence 4, Applicatio	1.31e-69
31	951	34.2	355	8	US-08-450-Sequence 8, Applicatio	1.31e-69
32	951	34.2	355	7	US-08-308-Sequence 4, Applicatio	1.31e-69
33	951	34.2	1065	10	US-08-605-Sequence 22, Applicati	1.31e-69
34	948	34.1	359	6	US-08-245-Sequence 24, Applicati	2.39e-69
35	927	33.4	360	12	US-08-887-Sequence 16, Applicati	1.63e-67
36	927	33.4	360	12	US-08-875-Sequence 20, Applicati	1.63e-67
37	927	33.4	360	14	US-09-067-Sequence 34, Applicati	1.63e-67
38	927	33.4	360	14	US-08-675-Sequence 14, Applicati	1.63e-67
39	927	33.4	360	14	US-09-045-Sequence 55, Applicati	1.63e-67
40	928	33.4	378	3	US-07-980-Sequence 2, Applicatio	1.33e-67
41	928	33.4	378	3	US-07-980-Sequence 2, Applicatio	1.33e-67
42	928	33.4	378	3	US-07-980-Sequence 2, Applicatio	1.33e-67
43	928	33.4	378	14	US-09-045-Sequence 49, Applicati	1.33e-67
44	928	33.4	378	3	US-07-980-Sequence 2, Applicatio	1.33e-67
45	928	33.4	378	7	US-08-352-Sequence 2, Applicatio	1.33e-67

ALIGNMENTS

RESULT 1  
ID US-08-675-814-8 STANDARD; PRT; 365 AA.  
XX  
AC  
XX  
XX  
XX  
XX  
DE Sequence 8, Application US/08675814  
XX  
CC Sequence 8, Application US/08675814  
CC GENERAL INFORMATION:  
CC APPLICANT: Gish, Kurt C.  
CC APPLICANT: Schall, Thomas J.  
CC APPLICANT: Vicari, Alain  
CC APPLICANT: Zlotnik, Albert  
CC TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS  
CC NUMBER OF SEQUENCES: 14  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: DNAX Research Institute  
CC STREET: 901 California Avenue  
CC CITY: Palo Alto  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94304-1104  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/675,814  
CC FILING DATE: 05-JUL-1996  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Ching, Edwin P.  
CC REGISTRATION NUMBER: 34,090  
CC REFERENCE/DOCKET NUMBER: DX0589  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 415-852-9196  
CC TELEFAX: 415-496-1200  
CC INFORMATION FOR SEQ ID NO: 8:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 365 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear



CC MOLECULE TYPE: protein  
SQ SEQUENCE 365 AA; 41375 MW; 721306 CN;  
Query Match 100.0%; Score 2779; DB 10; Length 365;  
Best Local Similarity 100.0%; Pred. No. 1.60e-231;  
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 MFSTPVKIIICQSLHITQILRCYCPCRRSGSPGYLYRIAYSLICVLGLGNILVVI 60  
QY 1 MFSTPVKIIICQSLHITQILRCYCPCRRSGSPGYLYRIAYSLICVLGLGNILVVI 60  
Db 61 TFAYFKARSMTDVYLLNMAIADILFVLTLPFWAVSHATGAWVFSNATCKLLKGIYAINF 120  
QY 61 TFAYFKARSMTDVYLLNMAIADILFVLTLPFWAVSHATGAWVFSNATCKLLKGIYAINF 120  
Db 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKSKIIICLVVWGLSVIISSTFFVNOKY 180  
QY 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKSKIIICLVVWGLSVIISSTFFVNOKY 180  
Db 181 NTQGSVCEPKYQTVSEPIRWKLLMGLLELFGFPILPMFMIFCYTFIVKTLVQAQNSKR 240  
QY 181 NTQGSVCEPKYQTVSEPIRWKLLMGLLELFGFPILPMFMIFCYTFIVKTLVQAQNSKR 240  
Db 241 HKAIRVIAVVLVFLACQIPHNMVLLVTAANLGMNRSQSEKLGITKTKTVEVLAFHLC 300  
QY 241 HKAIRVIAVVLVFLACQIPHNMVLLVTAANLGMNRSQSEKLGITKTKTVEVLAFHLC 300  
Db 301 CLNPVLYAFIGQKFRNFYFLKILKDLWCVRKRYKSSGFSFCAGRYSENISROTSETADNDNA 360  
QY 301 CLNPVLYAFIGQKFRNFYFLKILKDLWCVRKRYKSSGFSFCAGRYSENISROTSETADNDNA 360  
Db 361 SSFTM 365  
QY 361 SSFTM 365  
RESULT 2  
ID US-08-887-977-10 STANDARD; PRT; 365 AA.  
AC xxxxxx  
DT  
XX  
DE Sequence 10, Application US/08887977  
CC GENERAL INFORMATION:  
CC APPLICANT: Wang, Wei  
CC APPLICANT: Gish, Kurt C.  
CC APPLICANT: Schall, Thomas J.  
CC APPLICANT: Vicari, Alain P.  
CC APPLICANT: Zlotnik, Albert  
CC TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS  
CC NUMBER OF SEQUENCES: 19  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: DNAX Research Institute  
CC STREET: 901 California Avenue  
CC CITY: Palo Alto  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94304-1104  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/887,977  
CC FILING DATE: 03-JUL-1997  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/021,644  
CC FILING DATE: 05-JUL-1996  
CC

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/028,329  
CC FILING DATE: 11-OCT-1996  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Ching, Edwin P.  
CC REGISTRATION NUMBER: 34,090  
CC REFERENCE/DOCKET NUMBER: DX0589K1  
CC TELEPHONE: 650-852-9192  
CC TELEFAX: 650-496-1200  
CC INFORMATION FOR SEQ ID NO: 10:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 365 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 365 AA; 41375 MW; 721306 CN;  
Query Match 100.0%; Score 2779; DB 12; Length 365;  
Best Local Similarity 100.0%; Pred. No. 1.60e-231;  
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 MFSTPVKIIICQSLHITQILRCYCPCRRSGSPGYLYRIAYSLICVLGLGNILVVI 60  
QY 1 MFSTPVKIIICQSLHITQILRCYCPCRRSGSPGYLYRIAYSLICVLGLGNILVVI 60  
Db 61 TFAYFKARSMTDVYLLNMAIADILFVLTLPFWAVSHATGAWVFSNATCKLLKGIYAINF 120  
QY 61 TFAYFKARSMTDVYLLNMAIADILFVLTLPFWAVSHATGAWVFSNATCKLLKGIYAINF 120  
Db 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKSKIIICLVVWGLSVIISSTFFVNOKY 180  
QY 121 NCGMLLTICISMDRYIAIVQATKSFRLRSRTLPKSKIIICLVVWGLSVIISSTFFVNOKY 180  
Db 181 NTQGSVCEPKYQTVSEPIRWKLLMGLLELFGFPILPMFMIFCYTFIVKTLVQAQNSKR 240  
QY 181 NTQGSVCEPKYQTVSEPIRWKLLMGLLELFGFPILPMFMIFCYTFIVKTLVQAQNSKR 240  
Db 241 HKAIRVIAVVLVFLACQIPHNMVLLVTAANLGMNRSQSEKLGITKTKTVEVLAFHLC 300  
QY 241 HKAIRVIAVVLVFLACQIPHNMVLLVTAANLGMNRSQSEKLGITKTKTVEVLAFHLC 300  
Db 301 CLNPVLYAFIGQKFRNFYFLKILKDLWCVRKRYKSSGFSFCAGRYSENISROTSETADNDNA 360  
QY 301 CLNPVLYAFIGQKFRNFYFLKILKDLWCVRKRYKSSGFSFCAGRYSENISROTSETADNDNA 360  
Db 361 SSFTM 365  
QY 361 SSFTM 365  
RESULT 3  
ID US-09-045-583-48 STANDARD; PRT; 374 AA.  
AC xxxxxx  
DT  
XX  
DE Sequence 48, Application US/09045583  
CC GENERAL INFORMATION:  
CC APPLICANT: Graham, Gerard J. et al.  
CC TITLE OF INVENTION: Novel Molecules of the G Protein-Coupled  
CC NUMBER OF SEQUENCES: 56  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: LAHIVE & COCKFIELD, LLP  
CC STREET: 28 State Street  
CC CITY: Boston  
CC STATE: Massachusetts  
CC COUNTRY: USA  
CC ZIP: 02109  
CC COMPUTER READABLE FORM:  
CC



CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/045,583  
CC FILING DATE: 20-MAR-98  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mandragoras, Amy E.  
CC REGISTRATION NUMBER: 36,207  
CC REFERENCE/DOCKET NUMBER: MNI-044  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (617)227-7400  
CC TELEFAX: (617)742-4214  
CC INFORMATION FOR SEQ ID NO: 48:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 374 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC FRAGMENT TYPE: internal  
CC SEQUENCE 374 AA; 42494 MW; 759989 CN;

Query Match 88.0%; Score 2445; DB 14; Length 374;  
Best Local Similarity 98.8%; Pred. No. 9,38e-202;  
Matches 324; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Db 47 LFVPTAYSILICVGLLGNLIVITFAFYKARSMTDYLNNMAIADILFLVTLFPFVAVSH 106  
Qy 39 LY-RIAYSILICVGLLGNLIVITFAFYKARSMTDYLNNMAIADILFLVTLFPFVAVSH 97  
Db 107 ATGAWFVSNTCKLLKGIXAINFNCGMLLTICISMDRYIAIVQATKFRSLRSLRPSKI 166  
Qy 98 ATGAWFVSNTCKLLKGIXAINFNCGMLLTICISMDRYIAIVQATKFRSLRSLRPSKI 157  
Db 167 ICLVWGLSVIISSTFVNQKNTQGSVDCEPKYQTVSEPIRWKLLMLGLELLFGFFIP 226  
Qy 158 ICLVWGLSVIISSTFVNQKNTQGSVDCEPKYQTVSEPIRWKLLMLGLELLFGFFIP 217  
Db 227 LMFMTFCYTFIVKTLVQAKNSRHKRAIRVIAVVLFLACQIPHNMVLLVTAANLGRNR 286  
Qy 218 LMFMTFCYTFIVKTLVQAKNSRHKRAIRVIAVVLFLACQIPHNMVLLVTAANLGRNR 277  
Db 287 SCQSEKLGITVTVEVLAFLHCLNPVLYAFIGQKFRNFTFLKLDLWCVRKRYKSSGF 346  
Qy 278 SCQSEKLGITVTVEVLAFLHCLNPVLYAFIGQKFRNFTFLKLDLWCVRKRYKSSGF 337  
Db 347 SCAGRYSENISQTSQTSADNDNASSFTM 374  
Qy 338 SCAGRYSENISQTSQTSADNDNASSFTM 365

RESULT 4  
ID US-08-802-627A-2 STANDARD; PRT; 350 AA.

XX xxxxxx

AC xxxxxx

DT

DE Sequence 2, Application US/08802627A

XX Sequence 2, Application US/08802627A

CC GENERAL INFORMATION:

CC APPLICANT: Lee, James

CC APPLICANT: Wood, William I.

CC TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR

CC NUMBER OF SEQUENCES: 6

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Genentech, Inc.

CC STREET: 460 Point San Bruno Blvd  
CC CITY: South San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94080  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Winpatin (Genentech)  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/802,627A  
CC FILING DATE: 19-Feb-1997  
CC CLASSIFICATION: 514  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/284586  
CC FILING DATE: 10-AUG-1994  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/076093  
CC FILING DATE: 11-JUN-1993  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 07/810782  
CC FILING DATE: 19-DEC-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Love, Richard B.  
CC REGISTRATION NUMBER: 34,659  
CC REFERENCE/DOCKET NUMBER: P0706P2P1D2  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 415/225-5530  
CC TELEFAX: 415/953-9881  
CC TELEX: 910/371-7168  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 350 amino acids  
CC TYPE: Amino Acid  
CC TOPOLOGY: Linear  
CC SEQUENCE 350 AA; 39805 MW; 560082 CN;

Query Match 35.4%; Score 984; DB 12; Length 350;  
Best Local Similarity 44.4%; Pred. No. 1.72e-72;  
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETFELNKYVLIAYALVELLSLGNLSVLMVLISRVGRSVDYVLLNALAD 85  
Qy 25 YCAPCRSGSPG-YLYRIAYSLICVLGLLGNLIVITFAFYKARSMTDYLNNMAIAD 83  
Db 86 LLFALTLPWAASKVNG-WIFGTFLCKVVSLLKEVNSYSGILLACISVDYVLAIVHATR 144  
Qy 84 ILFVLTLPFVAVSHATGAWFVSNTCKLLKGIXAINFNCGMLLTICISMDRYIAIVQATK 143  
Db 145 TLQK-RHL--VKFVCLGCGWLSNMNLSLPPFLFRQAYHPNNSSPVC---YEVLGNDTAKW 198  
Qy 144 SFLRSRTPRSKIICLVWGLSVIISSTFVNQKNTQGS-DVCEPKYQTVS-EPFRW 201  
Db 199 RMVRLPHTFGFIVPLFVMLFCYGFTRTLFLKAHMGOKHRAVRVFAVLIIFLCWLPY 258  
Qy 202 KLLMLGLELLFGFFITPLFMIFCYTFIVKTLVQAKNSRHKRAIRVIAVVLFLACQIPH 261  
Db 259 NLVLLADTLMETQVIOETCERRNNIGRALDATEILGFLHSLCLNPIIYAFIGNFRHGLK 318  
Qy 262 NMVLLVTANLGRNR-SCQSEKLGITKTVTEVLAFLHCLNPVLYAFIGQKFRNFTFLK 320  
Db 319 IL 320  
Qy 321 IL 322

RESULT 5  
ID US-09-104-296-2 STANDARD; PRT; 350 AA.

XX xxxxxx

AC xxxxxx

DT



```
XX DE Sequence 2, Application US/09104296
XX CC Sequence 2, Application US/09104296
CC CC GENERAL INFORMATION:
CC CC APPLICANT: Lee, James
CC CC APPLICANT: Wood, William I.
CC CC TITLE OF INVENTION: PF4A Receptors
CC CC NUMBER OF SEQUENCES: 6
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: Genentech, Inc.
CC CC STREET: 1 DNA Way
CC CC CITY: South San Francisco
CC CC STATE: California
CC CC COUNTRY: USA
CC CC ZIP: 94080
CC CC COMPUTER READABLE FORM:
CC CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
CC CC COMPUTER: IBM PC compatible
CC CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC CC SOFTWARE: WinPatIn (Genentech)
CC CC CURRENT APPLICATION DATA:
CC CC APPLICATION NUMBER: US/09/104,296
CC CC FILING DATE: 24-June-1998
CC CC CLASSIFICATION:
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: 08/701265
CC CC FILING DATE: 22-AUG-1996
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: 08/664228
CC CC FILING DATE: 06-JUN-1996
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: 08/076093
CC CC FILING DATE: 11-JUN-1993
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: 07/810782
CC CC FILING DATE: 19-DEC-1991
CC CC ATTORNEY/AGENT INFORMATION:
CC CC NAME: Love, Richard B.
CC CC REGISTRATION NUMBER: 34,659
CC CC REFERENCE/DOCKET NUMBER: P0706P2C2
CC CC TELEPHONE: 415/225-5530
CC CC TELEFAX: 415/952-9881
CC CC TELEX: 910/371-7168
CC CC INFORMATION FOR SEQ ID NO: 2:
CC CC SEQUENCE CHARACTERISTICS:
CC CC LENGTH: 350 amino acids
CC CC TYPE: Amino Acid
CC CC TOPOLOGY: Linear
CC CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 15; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETLNKYVVIAYALVFLSLGNSLVMLVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRSGSSPG-YLYRIAYSLICVLGLGNILVITFAFYKKARSMVTDVYLLNMAIAD 83
Db 86 LFLALTLPWAASKVNG-WIFGTFLCKVSVLLKEVNFYSGILLACISVDYRLAIVHATR 144
QY 84 ILFVLTLFPWAVSHATGAWVFSNATCKLLKGIYAINFNCGMILLTLCISMDRYIAIVQATK 143
Db 145 TLTKQK-RHL--VKFVCLGCGWGLSMNLSLPFLFRQAYHPNNSPVC---YEVLGNDTAKW 198
QY 144 SFRLSRTPRSKIICLVVWGLSVIISSTFVFNQKNTQGS-DVCEPKYQTVS-EPIRW 201
Db 199 RAVLRILPHFGFTVPLFVLMFCYCGFTLRTLFKAHMGOKHRAVRVFAVILFICLWLPY 258
QY 202 KLLMLGLELFGFIPLFMFIFCYTFVTKLVQANSKRHKAIRVIAVILVFLACQIPH 261
Db 259 NUVLADTLMTQVTOETCERRNNIGRALDATEIGLFLHSLCNPPIYAFIGNGRHFLK 318

QY 262 NMVLLVTAANLKMNR-SCQSEKLGTYTIVTEVLAFLHCLLPVLYAFIGKFRNYFLK 320
Db 319 IL 320
QY 321 IL 322

RESULT 6
ID US-08-450-393-7 STANDARD; PRT; 350 AA.
XX AC XXXXXX
XX CC
XX CC
XX CC Sequence 7, Application US/08450393
XX CC Sequence 7, Application US/08450393
CC CC GENERAL INFORMATION:
CC CC APPLICANT: Charo, Israel
CC CC APPLICANT: Coughlin, Shaun
CC CC TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
CC CC TITLE OF INVENTION: PROTEIN RECEPTORS
CC CC NUMBER OF SEQUENCES: 14
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
CC CC STREET: 5 Palo Alto Square
CC CC CITY: Palo Alto
CC CC STATE: California
CC CC COUNTRY: USA
CC CC ZIP: 94308-2155
CC CC COMPUTER READABLE FORM:
CC CC MEDIUM TYPE: Floppy disk
CC CC COMPUTER: IBM PC compatible
CC CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CC CURRENT APPLICATION DATA:
CC CC APPLICATION NUMBER: US/08/450,393
CC CC FILING DATE: May 25, 1995
CC CC CLASSIFICATION: 424
CC CC ATTORNEY/AGENT INFORMATION:
CC CC NAME: Cseir, Luann
CC CC REGISTRATION NUMBER: 31,822
CC CC REFERENCE/DOCKET NUMBER: UCAL-237/02US
CC CC TELECOMMUNICATION INFORMATION:
CC CC TELEPHONE: 415-843-5165
CC CC TELEFAX: 415-8857-0663
CC CC TELEX: 380816CooleyPA
CC CC INFORMATION FOR SEQ ID NO: 7:
CC CC SEQUENCE CHARACTERISTICS:
CC CC LENGTH: 350 amino acids
CC CC TYPE: amino acid
CC CC TOPOLOGY: linear
CC CC MOLECULE TYPE: protein
CC CC HYPOTHETICAL: NO
CC CC ANTI-SENSE: NO
CC CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 8; Length 350;
Best Local Similarity 44.4%; Pred. No. 1.72e-72;
Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETLNKYVVIAYALVFLSLGNSLVMLVILYSRVGRSVTDVYLLNLALAD 85
QY 25 YCAPCRSGSSPG-YLYRIAYSLICVLGLGNILVITFAFYKKARSMVTDVYLLNMAIAD 83
Db 86 LFLALTLPWAASKVNG-WIFGTFLCKVSVLLKEVNFYSGILLACISVDYRLAIVHATR 144
QY 84 ILFVLTLFPWAVSHATGAWVFSNATCKLLKGIYAINFNCGMILLTLCISMDRYIAIVQATK 143
Db 145 TLTKQK-RHL--VKFVCLGCGWGLSMNLSLPFLFRQAYHPNNSPVC---YEVLGNDTAKW 198
QY 144 SFRLSRTPRSKIICLVVWGLSVIISSTFVFNQKNTQGS-DVCEPKYQTVS-EPIRW 201
```







CC TELEFAX: 415/952-9881  
 CC TELEX: 910/371-7168  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 350 amino acids  
 CC TYPE: Amino Acid  
 CC TOPOLOGY: Linear  
 CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 6; Length 350;  
 Best Local Similarity 44.4%; Pred. No. 1.72e-72;  
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETLNKYVVIYAIYALVFLSLGLNSVLMVILYSRVGRSVTDVYLLNLALAD 85  
 QY 25 YCAPCRSSGSPG-YLYRIAYSILICVLGLNLVITFAFYKKARSMTDVYLLNMAIAD 83  
 Db 86 LLFALTLPWAASKVNG-WIFGTFCKVWSLLKEVNFYSGILLACISVDRIYLAIVHATR 144  
 QY 84 ILFVLTLPFWAYSHATGANVFSNATCKLLGIYAINFCNCGMLLTCTISMDRYIAIVQATK 143  
 Db 145 TLTKQ-RHL--VKFVCLGCGWGLSNLSLPFFLFRQAYHPNNSPVC---YEVLGNDTAKW 198  
 QY 144 SFLRSRTLPKRSKIICLVVWGLSVIISSTFFVNQYNTQSD-VCEPKYQTVS-EPIRW 201  
 Db 199 RMVLRIILPHFTFGFIVPLFVLMFCYGFTRTLFRAHMGOKHRAMRVFAVVLIFLLCWLPY 258  
 QY 202 KLMLGLELLFGFFIPLFMFICYTFIVKTLVQAQNSKRHKRAIRVIAVVLFLACQIPH 261  
 Db 259 NLVLLADTLMTQVIOETCERRNIGRALDATEILGLHSCNLPYIAFYAGQNFHGFLEK 318  
 QY 262 NMVLLVTANLGMNR-SCQSEKLGITKTVTEVLAFLHCCNLNPLVIAFYAGQNFHGFLEK 320  
 Db 319 IL 320  
 QY 321 IL 322

RESULT 9  
 ID PCT-US93-10672-3 STANDARD; PRT; 350 AA.  
 XX  
 AC xxxxxx  
 DT  
 DE  
 XX

Sequence 3, Application PC/TUS9310672

Sequence 3, Application PC/TUS9310672  
 GENERAL INFORMATION:  
 CC APPLICANT: GENENTECH, INC.  
 CC TITLE OF INVENTION: CC-CHEMOKINE RECEPTOR  
 CC NUMBER OF SEQUENCES: 11  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Genentech, Inc.  
 CC STREET: 460 Point San Bruno Blvd  
 CC CITY: South San Francisco  
 CC STATE: California  
 CC COUNTRY: USA  
 CC ZIP: 94080  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: patin (Genentech)  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US93/10672  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER:  
 CC FILING DATE:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Fitts, Renee A.

CC REGISTRATION NUMBER: 35,136  
 CC REFERENCE/DOCKET NUMBER: 806  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415/225-1489  
 CC TELEFAX: 415/952-9881  
 CC TELEX: 910/371-7168

CC INFORMATION FOR SEQ ID NO: 3:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 350 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 1; Length 350;  
 Best Local Similarity 44.4%; Pred. No. 1.72e-72;  
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;

Db 27 Y-SPCMLETLNKYVVIYAIYALVFLSLGLNSVLMVILYSRVGRSVTDVYLLNLALAD 85  
 QY 25 YCAPCRSSGSPG-YLYRIAYSILICVLGLNLVITFAFYKKARSMTDVYLLNMAIAD 83  
 Db 86 LLFALTLPWAASKVNG-WIFGTFCKVWSLLKEVNFYSGILLACISVDRIYLAIVHATR 144  
 QY 84 ILFVLTLPFWAYSHATGANVFSNATCKLLGIYAINFCNCGMLLTCTISMDRYIAIVQATK 143  
 Db 145 TLTKQ-RHL--VKFVCLGCGWGLSNLSLPFFLFRQAYHPNNSPVC---YEVLGNDTAKW 198  
 QY 144 SFLRSRTLPKRSKIICLVVWGLSVIISSTFFVNQYNTQSD-VCEPKYQTVS-EPIRW 201  
 Db 199 RMVLRIILPHFTFGFIVPLFVLMFCYGFTRTLFRAHMGOKHRAMRVFAVVLIFLLCWLPY 258  
 QY 202 KLMLGLELLFGFFIPLFMFICYTFIVKTLVQAQNSKRHKRAIRVIAVVLFLACQIPH 261  
 Db 259 NLVLLADTLMTQVIOETCERRNIGRALDATEILGLHSCNLPYIAFYAGQNFHGFLEK 318  
 QY 262 NMVLLVTANLGMNR-SCQSEKLGITKTVTEVLAFLHCCNLNPLVIAFYAGQNFHGFLEK 320  
 Db 319 IL 320  
 QY 321 IL 322

RESULT 10  
 ID US-08-05-478-2 STANDARD; PRT; 350 AA.  
 XX  
 AC xxxxxx  
 DT  
 DE  
 XX

Sequence 2, Application US/08805478

Sequence 2, Application US/08805478  
 GENERAL INFORMATION:  
 CC APPLICANT: Chuntharapai, Anan  
 CC APPLICANT: Lee, James  
 CC APPLICANT: Hebert, Caroline  
 CC APPLICANT: Jin Kim, K.  
 CC TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR  
 CC NUMBER OF SEQUENCES: 6  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Genentech, Inc.  
 CC STREET: 460 Point San Bruno Blvd  
 CC CITY: South San Francisco  
 CC STATE: California  
 CC COUNTRY: USA  
 CC ZIP: 94080  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Winpatin (Genentech)  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/05,478







CC GENERAL INFORMATION:  
 CC APPLICANT: RICHARD HORUK, KULDEEP NEOTE, THOMAS SCHALL  
 CC TITLE OF INVENTION: CC-CHEMOKINE RECEPTOR  
 CC NUMBER OF SEQUENCES: 11  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Genentech, Inc.  
 CC STREET: 460 Point San Bruno Blvd  
 CC CITY: South San Francisco  
 CC STATE: California  
 CC COUNTRY: USA  
 CC ZIP: 94080  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: patin (Genentech)  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/308,880  
 CC FILING DATE:  
 CC CLASSIFICATION: 530  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 07/974025  
 CC FILING DATE: 10-NOV-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Fitts, Renee A. 35,136  
 CC REGISTRATION NUMBER:  
 CC REFERENCE/DOCKET NUMBER: 806  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415/225-1489  
 CC TELEFAX: 415/952-9881  
 CC TELEX: 910/371-7168  
 CC INFORMATION FOR SEQ ID NO: 3:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 350 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC SQ SEQUENCE 350 AA; 39805 MW; 660082 CN;

Query Match 35.4%; Score 984; DB 7; Length 350;  
 Best Local Similarity 44.4%; Pred. No. 1.72e-72;  
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;  
 Db 27 Y-SPCMLETFELNKVYVIAVALFLLSGNSLVMLVILYKSRGSRVTDVYLLNLALAD 85  
 QY 25 YCAPCRSSGSSPG-YLYRIAYSLICVLGLGNILVITFAFYKARSDTDVYLLNMAIAD 83  
 Db 86 LFLALTLPIWAASKVNG-WIFGTFLCKVSLKKEVNFYSGILLACISVDRIYAIHVATR 144  
 QY 84 ILFVLTLFPWAVSHATGAWFVSNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATK 143  
 Db 145 TLTKQ-RHL--VKFVCLGCGWLSMNLSPFFLFRQAYHPNNSPVC---YEVLGNDTAKW 198  
 QY 144 SFRLRSRTLPKRSKIICLVWGLSVIISSTFVFNQKYNTOGSD-VCEPKYQTVS-EPIRW 201  
 Db 199 RMVLRILPHTFGFIVPLFVLMFCYGTFTLTLEKHAHQKRAHMRVIFAVVLIPLCLWLPY 258  
 QY 202 KLLMLGLELFGFFPLMFIFCYFTIVKTLVQAQNSKRKRAIRVIAVVLVFLACQIPH 261  
 Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGLFSLCLNPIIYAFIQNFRHGFLK 318  
 QY 262 NMVLLVTAANLGMKMR-SCQSEKLIQYTKTVTEVLAFLHCLLPVLYAFIQGKFRNYFLK 320  
 Db 319 IL 320  
 QY 321 IL 322

RESULT 13  
 ID US-08-441-562-3 STANDARD; PRT; 350 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT

XX Sequence 3, Application US/08441562  
 DE  
 XX Sequence 3, Application US/08441562  
 CC GENERAL INFORMATION:  
 CC APPLICANT: RICHARD HORUK, KULDEEP NEOTE, THOMAS SCHALL  
 CC TITLE OF INVENTION: Nucleic Acids Encoding C-C Chemokine Receptor  
 CC NUMBER OF SEQUENCES: 11  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Genentech, Inc.  
 CC STREET: 460 Point San Bruno Blvd  
 CC CITY: South San Francisco  
 CC STATE: California  
 CC COUNTRY: USA  
 CC ZIP: 94080  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: WinPatIn (Genentech)  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/441,562  
 CC FILING DATE:  
 CC CLASSIFICATION: 536  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 08/308880  
 CC FILING DATE: 19-SEP-1994  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 07/974025  
 CC FILING DATE: 10-NOV-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Love, Richard B.  
 CC REGISTRATION NUMBER: 34,659  
 CC REFERENCE/DOCKET NUMBER: P0806C1D1  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 415/225-5530  
 CC TELEFAX: 415/952-9881  
 CC TELEX: 910/371-7168  
 CC INFORMATION FOR SEQ ID NO: 3:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 350 amino acids  
 CC TYPE: Amino Acid  
 CC TOPOLOGY: Linear  
 CC SQ SEQUENCE 350 AA; 39805 MW; 660082 CN;  
 Query Match 35.4%; Score 984; DB 8; Length 350;  
 Best Local Similarity 44.4%; Pred. No. 1.72e-72;  
 Matches 134; Conservative 76; Mismatches 80; Indels 12; Gaps 9;  
 Db 27 Y-SPCMLETFELNKVYVIAVALFLLSGNSLVMLVILYKSRGSRVTDVYLLNLALAD 85  
 QY 25 YCAPCRSSGSSPG-YLYRIAYSLICVLGLGNILVITFAFYKARSDTDVYLLNMAIAD 83  
 Db 86 LFLALTLPIWAASKVNG-WIFGTFLCKVSLKKEVNFYSGILLACISVDRIYAIHVATR 144  
 QY 84 ILFVLTLFPWAVSHATGAWFVSNATCKLLKGIYAINFCGMLLTICISMDRYIAIVQATK 143  
 Db 145 TLTKQ-RHL--VKFVCLGCGWLSMNLSPFFLFRQAYHPNNSPVC---YEVLGNDTAKW 198  
 QY 144 SFRLRSRTLPKRSKIICLVWGLSVIISSTFVFNQKYNTOGSD-VCEPKYQTVS-EPIRW 201  
 Db 199 RMVLRILPHTFGFIVPLFVLMFCYGTFTLTLEKHAHQKRAHMRVIFAVVLIPLCLWLPY 258  
 QY 202 KLLMLGLELFGFFPLMFIFCYFTIVKTLVQAQNSKRKRAIRVIAVVLVFLACQIPH 261  
 Db 259 NLVLLADTLMTQVIOETCERRNNIGRALDATEILGLFSLCLNPIIYAFIQNFRHGFLK 318  
 QY 262 NMVLLVTAANLGMKMR-SCQSEKLIQYTKTVTEVLAFLHCLLPVLYAFIQGKFRNYFLK 320  
 Db 319 IL 320  
 QY 321 IL 322



ID	US-08-801-238-2	STANDARD;	PRT;	350 AA.
Db	199	RMVLRILPHTGFI	VPLFVLMFLPCY	GFTLRTLFKAHQKRAHMRVIFAVVFLFLCWLXPY 25
Qy	202	KLLMLGLELLG	FFIPLMFIMFCY	TFIVKTLVQAQNSKSKHKAIRVIAVVLVFLACQIPH 261
Db	259	NVLVLADTLMR	TQVIQETCERN	NIIGRALDATEILGFLHSCNLPIIYATIGQNRHGFLK 318
Qy	262	NMVLVTAANLG	KMNR-SCQSEK	LIGYTKVTVEVLAFLHCLCLNPVLYAFYTGQKFRNYFLK 320
Db	319	IL	320	
Qy	321	IL	322	
RESULT	15			
ID	US-08-802-627-2	STANDARD;	PRT;	350 AA.
XX	XXXXXX			
DT	Sequence 2, Application US/08802627			
DE	Sequence 2, Application US/08802627			
CC	GENERAL INFORMATION:			
CC	APPLICANT: Lee, James			
CC	APPLICANT: Wood, William I.			
CC	TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR			
CC	NUMBER OF SEQUENCES: 6			
CC	CORRESPONDENCE ADDRESS:			
CC	ADDRESSEE: Genentech, Inc.			
CC	STREET: 460 Point San Bruno Blvd			
CC	CITY: South San Francisco			
CC	STATE: California			
CC	COUNTRY: USA			
CC	ZIP: 94080			
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk			
CC	COMPUTER: IBM PC compatible			
CC	OPERATING SYSTEM: PC-DOS/MS-DOS			
CC	SOFTWARE: WinPatIn (Genentech)			
CC	CURRENT APPLICATION DATA:			
CC	APPLICATION NUMBER: US/08/801,238			
CC	FILING DATE: 19-Feb-1997			
CC	CLASSIFICATION: 530			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: 08/284586			
CC	FILING DATE: 10-AUG-1994			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: 08/076093			
CC	FILING DATE: 11-JUN-1993			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: 07/810782			
CC	FILING DATE: 19-DEC-1991			
CC	ATTORNEY/AGENT INFORMATION:			
CC	NAME: Love, Richard B.			
CC	REGISTRATION NUMBER: 34,659			
CC	REFERENCE/DOCKET NUMBER: P0706P2PID1			
CC	TELECOMMUNICATION INFORMATION:			
CC	TELEPHONE: 415/225-5530			
CC	TELEFAX: 415/952-9881			
CC	TELEX: 910/371-7168			
CC	INFORMATION FOR SEQ ID NO: 2:			
CC	SEQUENCE CHARACTERISTICS:			
CC	LENGTH: 350 amino acids			
CC	TYPE: Amino Acid			
CC	TOPOLOGY: Linear			
CC	SEQUENCE 350 AA; 39805 MW; 660082 CN;			
Query Match	35.4%;	Score 984;	DB 12;	Length 350;
Best Local Similarity	44.4%;	Pred. No. 1.72e-72;		
Matches	134;	Conservative	76;	Mismatches 80; Indels 12; Gaps 9;
Db	27	Y-SPCMLETLNKYV	VIAYALVLLSGLNS	LMVLVLYSRVGRSVTDVYLLNLALAD 85
Qy	25	YCAPCRSSGSPG	-LYRIAYSLVCLG	LNLVITVITAFYKARSMTDVYLLNMAIAD 83
Db	86	LLFALTLPWAASK	VNG-WIFGTFCKV	YSLKEVNFYSGILLACISVDVRYIAVHATR 144
Qy	84	ILFVLTPFWAVSH	ATGAVVFNFSNATC	KLLKGIYAINFNCGMLLLTLCISMDRYIAIVQATK 143
Db	145	TLTOK-RHL--VK	FVCLCGWGLSMNLS	LPFLFRQAYHNNSSPVC---YEVLGNDTAKW 198
Qy	144	SFLRSRTPRKS	KIICLVVWVGLSV	ISSSTTFVFNQKINTQGSQD-VCEPKYQTVS-EPIRW 201



QY 25 YCAPCRRSSPG-YLYRTAYSLLICVLGILLGNILVITFAFYKKASMTDYYLLNMAID 83  
Db 86 LLFALTLPWAASKVNG-WIFGTFELCKVVSLLKEVNFYSGILLACISVDYRLAIYHATR 144  
QY 84 ILFVLTLPFWAYSHATGAWFVSNAACKLLKGIYAINFNCGMLLLTCTISMDRYIAIVQATK 143  
Db 145 TLTQK-RHL--VKFVCLGCGLSMNLSPFFLPRQAYHPNNSPVC---YEVLGNDTAKW 198  
QY 144 SPRLRSRTLPRKIIICLVVWGLSVIISSTFVNQKYNTOGSD-VCEPKYQTVS-EPIRW 201  
Db 199 RMVLRILPHTFGFIVPLFVMLFCYGFTRLTKAHMGOKHRAMRVIFAVVLIIFLLCWL PY 258  
QY 202 KLMLGLELLFGFFPLPMFMIICYTFIVTLVQAKSKRHKARVIAVVLVFLACQIPH 261  
Db 259 NLVLLADTLMRTQVTOETCERRNNIGRALDATEILGFLHSCLNPIIYAFYQNFHGFELK 318  
QY 262 NMVLLVTAANLGKMR-SCQSEKLIQYTKTVTEVLAFLHCCLNPVLYAFYQNFYELK 320  
Db 319 IL 320  
QY 321 IL 322

Search completed: Fri Nov 13 12:11:13 1998  
Job time : 132 secs.



!!SEQUENCE\_LIST 1.0  
(Nucleotide) WORDSEARCH of: /home/obryen/dra977/olig/us-08-887-977-9 check: 9278 from:  
FROMIG of: /home/obryen/dra977/olig/US08887977.seq  
sequence 9, application us/08887977  
general information:  
applicant: wang, wei  
applicant: gish, kurt c.  
applicant: schall, thomas j. . . .

TO: GenEMBL: \* Sequences: 567,306 Total-length: 1,118,886,981 November 16, 1998 15:09

Database Release Information:

GenBank, Release 109.0, Released on 15Oct1998, Formatted on 17Oct1998  
EMBL, Release 56.0, Released on 16Sep1998, Formatted on 18Oct1998

Word-size: 15 Words: 54754 Diagonals: 3,201 Total-diagonals: 2,000,000,000  
Integral-width: 1 Alphabet: 4 List-size: 50 CPU minutes: 118.46

*# of sequences containing motifs at least 15 nt long*

Sequence Strd Diag Score Width Documentation ..

GB_PR2:HSU45984	+	254	992	2	U45984 Homo sapiens CCR6 chemokine re
GB_PR1:HSCRL3	+	87	992	2	279784 H.sapiens G protein-coupled re
GB_PR3:HSU68032	+	222	990	2	U68032 Human G protein-coupled recept
GB_PR3:HSU68030	+	291	990	2	U68030 Human G protein-coupled recept
GB_PR1:HSU60000	+	27	930	2	U60000 Human IL8-related receptor (DR
GB_RO:AB009369	+	149	64	1	AB009369 Mus musculus mRNA for G prot
GB_IN:AC005148	+	894	19	1	AC005148 Drosophila melanogaster DNA
GB_IN:CEZK678	+	22866	19	1	279605 Caenorhabditis elegans cosmid
GB_HTG:AC004442	-	33655	19	1	AC004442 *** SEQUENCING IN PROGRESS *
GB_HTG:AC004311	-	29041	19	1	AC004311 *** SEQUENCING IN PROGRESS *
GB_RO:MAJ3959	-	1325	19	1	AJ223959 Mus musculus mRNA for very-l
GB_RO:AC072760	-	1156	19	1	AF072760 Mus musculus fatty acid tran
GB_PAT:112527	-	1714	19	1	112527 Sequence 1 from patent US 5428
GB_PAT:AB013985	-	-513	19	1	AB013985 Antirrhinum majus transposon
GB_SFS:G39250	+	-35	19	1	G39250 Z20386 zebrafish AB Dario reti
GB_PR2:HSBTDSS2	+	228	19	1	AF018631 Homo sapiens biotinidase (BTI
GB_PR2:HS93H18	+	69237	19	1	284488 Human DNA sequence from PAC 93
GB_PR2:AC002542	-	136361	19	1	AC002542 Human BAC clone RG114A06 fr
GB_HTG:HS347H13	-	8019	19	1	AL023553 Human DNA sequence *** SEQUE
GB_HTG:CEY7A5	+	116418	19	1	AL021576 Caenorhabditis elegans DNA
GB_IN:CFE26D10	+	17744	18	1	280223 Caenorhabditis elegans cosmid
GB_IN:CEC24H11	+	20715	18	1	281475 Caenorhabditis elegans cosmid
GB_IN:CEC24H11	+	24350	18	1	281475 Caenorhabditis elegans cosmid
GB_IN:AC004369	-	40361	18	1	AC004369 Drosophila melanogaster DNA
GB_IN:DSU63854	+	1152	18	1	U63857 Drosophila simulans decapenta
GB_IN:DMU63857	+	13516	18	1	U63857 Drosophila melanogaster decap
GB_IN:CEI24A11	+	27806	18	1	249072 Caenorhabditis elegans cosmid
GB_IN:CELT20F5	+	22293	18	1	U80442 Caenorhabditis elegans cosmid
GB_IN:CEUHSF70	-	917	18	1	M18540 C.elegans heat shock protein 7
GB_IN:CEH31B20	+	9079	18	1	AL021066 Caenorhabditis elegans cosmi
GB_HTG:AC004556	-	92474	18	1	AC004556 *** SEQUENCING IN PROGRESS *
GB_HTG:AC004479	-	176462	18	1	AC004479 *** SEQUENCING IN PROGRESS
GB_VI:CYCP	-	1647	18	1	X76716 Caliciviridae mRNA for capsid
GB_VI:CRNAORFS	-	5321	18	1	X86557 Human enteric calicivirus ORF1
GB_RO:AB009369	+	152	18	1	AB009369 Mus musculus mRNA for G prot
GB_PR3:HSU85195	+	48083	18	1	U85195 Homo sapiens BAC129, complete
GB_PR3:HSB107C2	+	7446	18	1	AL008724 Human genomic DNA sequence f
GB_PR3:HS211A9	+	50916	18	1	295889 Human DNA sequence from clone
GB_PR3:AC004993	+	96033	18	1	AC004993 Homo sapiens PAC clone DJ119
GB_PL1:PYTRBCLH	+	757	18	1	L34817 Pyramonas octopus chloroplast
GB_PL1:PYTRBCLG	+	757	18	1	L34816 Pyramonas parkeae chloroplast
GB_PL1:PYTRBCLAD	+	757	18	1	L34779 Pyramonas grossii chloroplast
GB_PL1:PYTRBCLAB	+	757	18	1	L34777 Pyramonas propulsa chloroplast
GB_PL1:PYTCPRBCLZ	+	757	18	1	L34834 Pyramonas sp. chloroplast r
GB_PL1:PEASHMTA	+	1217	18	1	M87649 Pisum sativum serine hydroxyme
GB_PL1:ATU20624	+	674	18	1	U20624 Arabidopsis thaliana calcium-b
GB_BA1:U67522	+	5501	18	1	U67522 Methanococcus jannaschii secti
GB_PR2:HUAE000658	+	48083	18	1	AE000658 Homo sapiens T-cell recepto
GB_PR2:HS125H2	+	39027	18	1	298949 Human DNA sequence from clone
GB_PR2:AC002511	+	29886	18	1	AC002511 Human DNA from chromosome 19



```

Release 3.1a John F. Collins/Bioinformatics Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

P3Parch.nm n.a. = n.a. database search, using Smith-Waterman algorithm
Run on: Sat Nov 14 01:13:49 1998; Map/Parch time 1495.31 Seconds
1539.698 Million cell updates/sec
Tabular output not generated.

```

```
>US-08-887-977-9
Description:
  (1-1119) from US0887977.seq
  1119
A.A. Sequence:
  1 ATGTTTGGACTCCAGTGA.....AGAAAGCTGAGTCTCCTAA 1119
  TACAAAGCTCAGAGGATCTT.....TCTTTGACATCGAGGATTT
```

```

Scoring table:  TABLE jmetric
Cap 10
Smatch STD :  Dbase 0; Query 0
Searched:      552174 seqs, 1021863385 bases x 2
Post-processing: Minimum Match 0%
                  Minimum first 1000 summaries

```

```
Database:
1:em ba 3:em.btg 3:em.bum1 4:em.bum2 5:em.in 6:em.com
7:em.or 8:em.ov 9:em.pat 10:em.pl 11:em.ro
genbank107

Database:
12:gb.ba 13:gb.btg 14:gb.in 15:gb.com 16:gb.ov 17:gb.pat
18:gb.pa 19:gb.pl 20:gb.prl 21:gb.pzr 22:gb.st 23:gb.st
24:gb.sta 25:gb.sty 26:gb.un 27:gb.v1

Mean 0.183: Variance 2.723: scale 3.372
Statistics:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

45	18	1.6	131387.13	H52917144	Human DNA sequence ***	1.07e+01
46	18	1.6	153117.13	AC004993	Human DNA sequence ***	1.07e+01
47	18	1.6	175353.21	H521582	Human DNA sequence ***	1.07e+01
48	18	1.6	180163.11	H521194	Human DNA sequence ***	1.07e+01
49	18	1.6	200109.13	H521582	Human DNA sequence ***	1.07e+01
50	18	1.6	250109.13	H521582	Human DNA sequence ***	1.07e+01
51	18	1.6	250529.13	H521582	Human DNA sequence ***	1.07e+01
52	18	1.6	260927.13	CE114C4	Caenorhabditis elegans	1.07e+01
53	18	1.6	260927.13	CE114C4	Caenorhabditis elegans	1.07e+01
54	18	1.6	260927.13	CE114C4	Caenorhabditis elegans	1.07e+01
55	17	1.5	430672.12	CE114C4	Caenorhabditis elegans	1.07e+01
56	17	1.5	430672.12	CE114C4	Caenorhabditis elegans	1.07e+01
57	17	1.5	172.20	H52917144	Human DNA sequence ***	1.07e+01
58	17	1.5	340.24	G35559	Human DNA sequence ***	1.07e+01
59	17	1.5	340.24	G35559	Human DNA sequence ***	1.07e+01
60	17	1.5	349.20	G03235	Human DNA sequence ***	1.07e+01
61	17	1.5	335.15	BTMS1875	B. taurus microRNA	1.07e+01
62	17	1.5	400.22	MR185	Mouse MHC class II	1.07e+01
63	17	1.5	400.22	MR185	Mouse MHC class II	1.07e+01
64	17	1.5	480.24	H52917144	Human DNA sequence ***	1.07e+01
65	17	1.5	536.22	H52917144	Human DNA sequence ***	1.07e+01
66	17	1.5	536.22	H52917144	Human DNA sequence ***	1.07e+01
67	17	1.5	536.22	H52917144	Human DNA sequence ***	1.07e+01
68	17	1.5	536.22	H52917144	Human DNA sequence ***	1.07e+01
69	17	1.5	738.19	STH130482	A. thaliana mRNA for RNA	1.07e+01
70	17	1.5	897.14	TEBVS678	T. brucei tandem telomere	1.07e+01
71	17	1.5	897.14	TEBVS678	T. brucei tandem telomere	1.07e+01
72	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
73	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
74	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
75	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
76	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
77	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
78	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
79	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
80	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
81	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
82	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
83	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
84	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
85	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
86	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
87	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
88	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
89	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
90	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
91	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
92	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
93	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
94	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01
95	17	1.5	1037.20	H52917144	Human DNA sequence ***	1.07e+01

U 08:55:24 1998

length of match

US-08-887-577-9.rye

and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	DB	ID	Description	Pred. No.
914	83.5	1253	21	HSRREG3	H. sapiens G protein-co	0.00e+00
914	83.5	1253	21	HSRREG3	H. sapiens G protein-co	0.00e+00
872	77.9	1518	21	HSU68032	Human G protein-couple	0.00e+00
872	77.9	1978	21	HSU68030	Human G protein-couple	0.00e+00

507 - new

## SUMMARIES

No.	Score	Match	Length	ID	Description	Pred.	No.
1	934	83.5	1255	21	HEKRL3	H. septena C protein-co	0.00e+00
2	932	83.5	1255	21	HEKRL3	H. septena C protein-co	0.00e+00
3	874	77.9	1518	21	HSU6R032	Human C protein-comple	0.00e+00
4	872	77.9	1518	21	HSU6R032	Human C protein-comple	0.00e+00
5	617	55.4	978	21	HSU6R030	Human C protein-comple	0.00e+00
6	1117	20	1117	20	HSU6R030	Human C protein-comple	0.00e+00
7	19	1.7	19	1.7	G32756	220R8 Zebrin/AB	1.77e+00
8	19	1.7	19	1.7	G32756	220R8 Zebrin/AB	1.77e+00
9	18	1.7	2248	19	AB017965	Antirrhinum majus zeta	1.77e+00
10	18	1.7	2248	19	AB017965	Antirrhinum majus zeta	1.77e+00
11	19	1.7	2310	22	MMJ31959	Mus musculus mRNA for	1.77e+00
12	19	1.7	2310	22	MMJ31959	Mus musculus mRNA for	1.77e+00
13	19	1.7	3253	17	1J1527	Sequence 1 from patent	1.77e+00
14	19	1.7	3253	17	1J1527	Sequence 1 from patent	1.77e+00
15	19	1.7	33124	14	CEK678	Cnemidophorus b. b. b.	1.77e+00
16	19	1.7	33124	14	CEK678	Cnemidophorus b. b. b.	1.77e+00
17	19	1.7	83942	13	AC004311	*** SEQUENCING IN PROG	1.77e+00
18	19	1.7	83942	13	AC004311	*** SEQUENCING IN PROG	1.77e+00
19	14	1.9	90527	13	CE7745	Cnemidophorus elegans	1.77e+00
20	14	1.9	90527	13	CE7745	Cnemidophorus elegans	1.77e+00
21	16	1.7	12425	14	AC003149	Drosophila melanogaster	1.77e+00
22	16	1.7	12425	14	AC003149	Drosophila melanogaster	1.77e+00
23	17	1.7	12425	14	AC003149	Drosophila melanogaster	1.77e+00
24	17	1.7	12425	14	AC003149	Drosophila melanogaster	1.77e+00
25	19	1.7	163741	21	BS39H18	Human BAC sequence for	1.77e+00
26	19	1.7	163741	21	BS39H18	Human BAC sequence for	1.77e+00
27	18	1.6	18078	21	AC002542	Human DNA sequence	1.77e+00
28	18	1.6	18078	21	AC002542	Human DNA sequence	1.77e+00
29	19	1.7	193118	13	HS347H13	Pyraminomonas sp. chloro	1.07e+01
30	19	1.7	193118	13	HS347H13	Pyraminomonas sp. chloro	1.07e+01
31	18	1.6	1089	19	PTPSC162	Pyraminomonas parvula c	1.07e+01
32	18	1.6	1089	19	PTPSC162	Pyraminomonas parvula c	1.07e+01
33	18	1.6	1089	19	PTPSC162	Pyraminomonas parvula c	1.07e+01
34	18	1.6	1089	19	PTPSC162	Pyraminomonas parvula c	1.07e+01
35	18	1.6	1089	19	PTPSC162	Pyraminomonas parvula c	1.07e+01
36	18	1.6	1803	19	PSASHTA	Pyraminomonas parvula c	1.07e+01
37	18	1.6	1803	19	PSASHTA	Pyraminomonas parvula c	1.07e+01
38	18	1.6	1896	19	ATU70624	Pisum sativum serine h	1.07e+01
39	18	1.6	1896	19	ATU70624	Pisum sativum serine h	1.07e+01
40	18	1.6	2327	14	CEK69770	Arabis thaliana c	1.07e+01
41	18	1.6	2327	14	CEK69770	Arabis thaliana c	1.07e+01
42	18	1.6	4298	14	CEK69770	C. elegans heat shock p	1.07e+01
43	18	1.6	4298	14	CEK69770	C. elegans heat shock p	1.07e+01
44	18	1.6	6935	21	AB002314	Caldicellulosymonas sp.	1.07e+01
45	18	1.6	6935	21	AB002314	Caldicellulosymonas sp.	1.07e+01
46	18	1.6	7555	27	CRM00RFS	Human enteric calliculic	1.07e+01
47	18	1.6	7555	27	CRM00RFS	Human enteric calliculic	1.07e+01
48	18	1.6	12890	14	CEK31820	Cnemidophorus funneus	1.07e+01
49	18	1.6	12890	14	CEK31820	Cnemidophorus funneus	1.07e+01
50	18	1.6	19436	14	CEK31820	Cnemidophorus funneus	1.07e+01
51	18	1.6	19436	14	CEK31820	Cnemidophorus funneus	1.07e+01
52	18	1.6	19436	14	CEK31820	Cnemidophorus funneus	1.07e+01
53	18	1.6	20065	14	CEK26810	Cnemidophorus elegans	1.07e+01
54	18	1.6	20065	14	CEK26810	Cnemidophorus elegans	1.07e+01
55	18	1.6	2600	14	CEK720F5	Cnemidophorus elegans	1.07e+01
56	18	1.6	2600	14	CEK720F5	Cnemidophorus elegans	1.07e+01
57	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
58	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
59	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
60	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
61	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
62	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
63	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
64	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
65	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
66	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
67	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
68	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
69	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
70	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
71	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
72	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
73	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
74	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
75	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
76	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
77	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
78	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
79	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
80	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
81	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
82	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
83	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
84	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
85	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
86	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
87	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
88	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
89	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
90	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
91	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
92	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
93	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
94	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
95	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
96	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
97	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
98	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
99	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01
100	18	1.6	3755	14	CEK724H1	Cnemidophorus elegans	1.07e+01

96	17	1.5	1060	21	AF013222	Homo sapiens isolate T	5.90e+01
97	17	1.5	1060	21	AF013223	Homo sapiens isolate T	5.90e+01
98	17	1.5	1060	21	AF013224	Homo sapiens isolate T	5.90e+01
99	17	1.5	1060	21	AF013505	Homo sapiens isolate T	5.90e+01
100	17	1.5	1060	21	AF013533	Homo sapiens isolate T	5.90e+01
101	17	1.5	1060	21	AF013534	Homo sapiens isolate T	5.90e+01
102	17	1.5	1060	21	AF013533	Homo sapiens isolate T	5.90e+01
103	17	1.5	1060	21	AF013501	Homo sapiens isolate T	5.90e+01
104	17	1.5	1060	21	AF013548	Homo sapiens isolate T	5.90e+01
105	17	1.5	1060	21	AF013548	Homo sapiens isolate T	5.90e+01
106	17	1.5	1060	21	AF013506	Homo sapiens isolate T	5.90e+01
107	17	1.5	1060	21	AF013521	Homo sapiens isolate T	5.90e+01
108	17	1.5	1060	21	AF013542	Pan troglodytes isolat	5.90e+01
109	17	1.5	1060	21	AF013507	Homo sapiens isolate T	5.90e+01
110	17	1.5	1060	21	AF013537	Homo sapiens isolate T	5.90e+01
111	17	1.5	1060	21	AF013536	Homo sapiens isolate T	5.90e+01
112	17	1.5	1060	21	AF013537	Homo sapiens isolate T	5.90e+01
113	17	1.5	1060	21	AF013500	Homo sapiens isolate H	5.90e+01
114	17	1.5	1060	21	AF013537	Homo sapiens isolate T	5.90e+01
115	17	1.5	1060	21	AF013539	Pan troglodytes isolat	5.90e+01
116	17	1.5	1060	21	AF013536	Homo sapiens isolate T	5.90e+01
117	17	1.5	1060	21	AF013536	Homo sapiens isolate T	5.90e+01
118	17	1.5	1060	21	AF013536	Homo sapiens isolate T	5.90e+01
119	17	1.5	1060	21	AF013539	Homo sapiens isolate T	5.90e+01
120	17	1.5	1060	21	AF013539	Homo sapiens isolate T	5.90e+01
121	17	1.5	1060	21	AF013548	Homo sapiens isolate T	5.90e+01
122	17	1.5	1060	21	AF013509	Homo sapiens isolate K	5.90e+01
123	17	1.5	1060	21	AF013537	Homo sapiens isolate T	5.90e+01
124	17	1.5	1060	21	AF013534	Homo sapiens isolate M	5.90e+01
125	17	1.5	1060	21	AF013538	Pan troglodytes isolat	5.90e+01
126	17	1.5	1060	21	AF013528	Homo sapiens isolate U	5.90e+01
127	17	1.5	1060	21	AF013528	Homo sapiens isolate U	5.90e+01
128	17	1.5	1060	21	AF013511	Homo sapiens isolate K	5.90e+01
129	17	1.5	1060	21	AF013512	Homo sapiens isolate M	5.90e+01
130	17	1.5	1060	21	AF013510	Homo sapiens isolate K	5.90e+01
131	17	1.5	1060	21	AF013510	Homo sapiens isolate K	5.90e+01
132	17	1.5	1060	21	AF013510	Homo sapiens isolate K	5.90e+01
133	17	1.5	1094	21	PF089377	Pomoe sapientis isolate R	5.90e+01
134	17	1.5	1098	21	CF0893374	Crocodylus acethops	5.90e+01
135	17	1.5	1086	21	CB0893324	Ruman CC chemokine rec	5.90e+01
136	17	1.5	1086	21	CB0893325	Crocodylus acethops	5.90e+01
137	17	1.5	1086	21	CB0893326	Pyraminomonas citranala	5.90e+01
138	17	1.5	1099	19	PF97666A	Pan troglodytes CC-che	5.90e+01
139	17	1.5	1147	12	HS087441	E. coli DNA polymerase	5.90e+01
140	17	1.5	1147	12	HS087442	Ruman CC heavy-chain I	5.90e+01
141	17	1.5	1220	30	HS087443	Human CC chemokine rec	5.90e+01
142	17	1.5	1225	21	HS057640	Zetrichella coli (clon	5.90e+01
143	17	1.5	1261	12	CB082E7A	Gallus gallus inward I	5.90e+01
144	17	1.5	1284	15	AF021141	Homo sapiens isolate R	5.90e+01
145	17	1.5	1376	21	HS053546	H. sapiens chemokine	5.90e+01
146	17	1.5	1376	21	HS053546	H. sapiens chemokine	5.90e+01



US-08-887-977-9. Page

C	147	17	1.5	1439	72	MUSGFPBP1	M.musculus mRNA for an
C	148	17	1.5	1439	72	MUSGFPBP1	muscle cytochrome P-450
C	149	17	1.5	1526	17	ELI5BP7	CNA encoding P450 1a
C	150	17	1.5	1542	22	RATCYB	Rat cytochrome P450 1a
C	151	17	1.5	1550	40	HMMJUGEN	Human Jpg and Jgm gene
C	152	17	1.5	1559	20	SH2B2	Human SH2B2
C	153	17	1.5	1569	32	SH2B3	Human SH2B3
C	154	17	1.5	1590	32	RATCYB2C	Rat cytochrome P-450 1
C	155	17	1.5	1747	16	AF002039	Xenopus laevis p21-act
C	156	17	1.5	1866	21	MM077672	Macaca mulatta CC chem
C	157	17	1.5	2001	30	SHMEL03	Human melanocyte prote
C	158	17	1.5	2049	32	SHMEL04	Human melanocyte prote
C	159	17	1.5	2049	32	MUSPHEA	Mouse HSC class I H2-K
C	160	17	1.5	2065	32	MUSPHEA	Mouse HSC class I H2-K
C	161	17	1.5	2103	19	SDRCL1A3	S.cerevisiae RCL11a ge
C	162	17	1.5	2107	19	DIADCLT03	Dicyostelium dimorphic
C	163	17	1.5	2107	19	DIADCLT03	Dicyostelium dimorphic
C	164	17	1.5	2210	30	HMMACH92	Human gamma tubulin 2
C	165	17	1.5	2210	30	HUNS61283	human chromosome 7 SFS
C	166	17	1.5	2377	22	RATP45J	Rat cytochrome P450C9
C	167	17	1.5	2411	16	CHRTKA	Chicken thymidine kina
C	168	17	1.5	2411	16	CHRTKA	Chicken thymidine kina
C	169	17	1.5	2543	20	HSICR3B3	H.sapiens immunoglobul
C	170	17	1.5	2543	20	HSICR3B3	H.sapiens immunoglobul
C	171	17	1.5	2543	20	HSICR3C	H.sapiens immunoglobul
C	172	17	1.5	2543	20	HSICR3C	H.sapiens immunoglobul
C	173	17	1.5	2543	20	HSICR3C	H.sapiens immunoglobul
C	174	17	1.5	2755	14	CEU61012	C.eubacterium thymidyl
C	175	17	1.5	2755	14	CEU61012	C.eubacterium thymidyl
C	176	17	1.5	3182	21	HUMGTGB1	Human Ig gamma J-mu
C	177	17	1.5	3689	21	HUMGTGB1	Human Ig gamma J-mu
C	178	17	1.5	3689	21	HUMGTGB1	Human Ig gamma J-mu
C	179	17	1.5	3423	17	CEU51974	C.eubacterium thymidyl
C	180	17	1.5	3423	17	D76444	Human hkf-1 mRNA, comp
C	181	17	1.5	3526	17	ELI3586	CNA encoding Rf1 pro
C	182	17	1.5	4072	22	MUS02038	Mus musculus G7ase-1c
C	183	17	1.5	4072	22	MUS02038	Mus musculus G7ase-1c
C	184	17	1.5	5271	22	MUS02034	Mouse H2C class I H2-K
C	185	17	1.5	5475	22	MM2AK2	Mouse H2-K(k) gene of
C	186	17	1.5	5593	22	MUS2H2C	Mouse H2C class I P5 g
C	187	17	1.5	6031	23	HSICR7AIG	H.sapiens DNA for BCL7
C	188	17	1.5	6031	23	HSICR7AIG	H.sapiens DNA for BCL7
C	189	17	1.5	6532	16	SHSCR5B2	Mouse sapiens CC chemok
C	190	17	1.5	6581	16	SHSCR5B2	Mouse sapiens CC chemok
C	191	17	1.5	7353	20	HSANKY	S.salar gene for growth
C	192	17	1.5	7423	21	AF002962	Human mRNA for ankyrin
C	193	17	1.5	7423	21	AF002962	Human mRNA for ankyrin
C	194	17	1.5	9051	10	SHM09TA01	Human erythroid alpha-
C	195	17	1.5	9051	10	SHM09TA01	Human erythroid alpha-
C	196	17	1.5	11240	12	CEU009864	Methanobacterium thym
C	197	17	1.5	12663	12	CEU009864	Methanobacterium thym
C	198	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	199	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	200	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	201	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	202	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	203	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	204	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	205	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	206	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	207	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	208	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	209	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	210	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	211	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	212	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	213	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	214	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	215	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	216	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	217	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	218	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	219	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	220	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	221	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	222	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	223	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	224	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	225	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	226	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	227	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	228	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	229	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	230	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	231	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	232	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	233	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	234	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	235	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	236	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	237	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	238	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	239	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	240	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	241	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	242	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	243	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	244	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	245	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	246	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	247	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	248	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	249	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	250	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	251	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	252	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	253	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	254	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	255	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	256	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	257	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	258	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	259	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	260	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	261	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	262	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	263	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	264	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	265	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	266	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	267	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	268	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	269	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	270	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	271	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	272	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	273	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	274	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	275	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	276	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	277	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	278	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	279	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	280	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	281	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	282	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	283	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	284	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	285	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	286	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	287	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	288	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	289	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	290	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	291	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	292	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	293	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	294	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	295	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	296	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	297	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	298	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	299	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	300	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	301	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	302	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	303	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	304	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	305	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	306	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	307	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	308	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	309	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	310	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	311	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	312	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	313	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	314	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	315	17	1.5	14160	12	CEU009864	Methanobacterium thym
C	316	17	1.5	14160	12	CEU009864</	

US-08-887-977-9. Iqo

c	198	17	1.5	16996	30	HS12P12B	Human DNA sequence frc	5.90e+01
c	199	17	1.5	17243	30	HS21419BC	Human DNA sequence frc	5.90e+01
c	200	17	1.5	17430	30	HS225010	Human DNA sequence frc	5.90e+01
c	201	17	1.5	22514	30	CE250C10	Cenorhabditis elegans	5.90e+01
c	202	17	1.5	24688	30	CE2F15C11	Cenorhabditis elegans	5.90e+01
c	203	17	1.5	24688	30	CE2F15C11	Cenorhabditis elegans	5.90e+01
c	204	17	1.5	26574	30	HS118E7	Cenorhabditis elegans	5.90e+01
c	205	17	1.5	27551	30	CE2F2C8	Cenorhabditis elegans	5.90e+01
c	206	17	1.5	27551	30	CE2F2C8	Cenorhabditis elegans	5.90e+01
c	207	17	1.5	30382	30	CE2F2A11	Cenorhabditis elegans	5.90e+01
c	208	17	1.5	31186	34	CE2C12A9	Cenorhabditis elegans	5.90e+01
c	209	17	1.5	33960	32	HABNC62	Mesorhizoctia auraria b	5.90e+01
c	210	17	1.5	34322	30	HS1917H3	Human DNA sequence frc	5.90e+01
c	211	17	1.5	34322	30	HS1917H3	Human DNA sequence frc	5.90e+01
c	212	17	1.5	35263	34	CE2F3A8	Cenorhabditis elegans	5.90e+01
c	213	17	1.5	36451	31	HS7021	Human (lambda) DNA for	5.90e+01
c	214	17	1.5	36565	31	CE2F31648	Human chromosome 11 72	5.90e+01
c	215	17	1.5	36649	30	HS190198	Human DNA sequence frc	5.90e+01
c	216	17	1.5	36693	34	CE2F10H9	Cenorhabditis elegans	5.90e+01
c	217	17	1.5	37773	31	AP027207	Homo sapiens chromosome 11	5.90e+01
c	218	17	1.5	38012	19	SP6C2P12	S. pombe chromosome II	5.90e+01
c	219	17	1.5	38425	21	HS191830	Human cosmid L121NC01	5.90e+01
c	220	17	1.5	40588	31	CE2F3A8	Cenorhabditis elegans	5.90e+01
c	221	17	1.5	40598	31	AC005306	Homo sapiens chromosome 11	5.90e+01
c	222	17	1.5	41599	13	AC004196	*** SQUZEMKING in PROG	5.90e+01
c	223	17	1.5	43952	21	HS101064	Human DNA sequence frc	5.90e+01
c	224	17	1.5	51409	13	HS10C11	Human DNA sequence **	5.90e+01
c	225	17	1.5	52724	13	AC002555	*** SQUZEMKING in PROG	5.90e+01
c	226	17	1.5	52724	13	AC002555	*** SQUZEMKING in PROG	5.90e+01
c	227	17	1.5	59060	31	AC005196	Homo sapiens chromosome 11	5.90e+01
c	228	17	1.5	73775	21	AP000026	Homo sapiens genomic	5.90e+01
c	229	17	1.5	79214	21	AC005159	Homo sapiens BAC clone	5.90e+01
c	230	17	1.5	80716	19	CE2F2A11	Cenorhabditis elegans	5.90e+01
c	231	17	1.5	80716	19	AP072897	Arabisopsis thaliana B	5.90e+01
c	232	17	1.5	84537	13	HS08081	Human DNA sequence **	5.90e+01
c	233	17	1.5	85535	19	ATAC0021388	Arabidopsis thaliana C	5.90e+01
c	234	17	1.5	88846	19	Z42424	Arabidopsis thaliana B	5.90e+01
c	235	17	1.5	92568	13	HS06D63	Human DNA sequence frc	5.90e+01
c	236	17	1.5	93695	19	AP071209	Arabisopsis thaliana D	5.90e+01
c	237	17	1.5	95643	19	F6N15	Arabidopsis thaliana D	5.90e+01
c	238	17	1.5	96636	13	AC005056	Homo sapiens clone RoD	5.90e+01
c	239	17	1.5	10113	13	CE2F3A8	Cenorhabditis elegans	5.90e+01
c	240	17	1.5	10113	13	AC007079	Cenorhabditis elegans	5.90e+01
c	241	17	1.5	102972	13	AC000360	*** SQUZEMKING in PROG	5.90e+01
c	242	17	1.5	110000	13	CE2F39A1	Cenorhabditis elegans	5.90e+01
c	243	17	1.5	110079	22	AC002135	Mouse BAC 146821 chrom	5.90e+01
c	244	17	1.5	110870	21	HS144724	Human DNA sequence frc	5.90e+01
c	245	17	1.5	118016	13	HS75399	Human DNA sequence **	5.90e+01
c	246	17	1.5	120000	21	AC003093	Human BAC clone RC385f	5.90e+01

18-08-887-977-9.198

249	17	1.5	122422	13	HS3A1D10	Human DNA sequence ***	5.90e+01
250	17	1.5	126885	13	AC004315	Human sapiens clone D10	5.90e+01
251	17	1.5	126886	13	AC004316	Human sapiens clone D10	5.90e+01
252	17	1.5	126887	13	AC004755	*** SEQUENCING IN PROG	5.90e+01
253	17	1.5	126894	21	AC003070	Human sapiens chromosome	5.90e+01
254	17	1.5	111000	13	HA46681	Human DNA sequence ***	5.90e+01
255	17	1.5	111001	13	HA46682	Human DNA sequence ***	5.90e+01
256	17	1.5	111523	21	HS04030	Human sapiens DNA seq	5.90e+01
257	17	1.5	111528	21	HS43014	Human sapiens DNA seq	5.90e+01
258	17	1.5	116742	12	EC002598	Escherichia coli genom	5.90e+01
259	17	1.5	119049	21	AC004147	Human sapiens clone D10	5.90e+01
260	17	1.5	120525	11	HS055226	Human sapiens frag (cc	5.90e+01
261	17	1.5	121268	11	HS055226	Human sapiens frag (cc	5.90e+01
262	17	1.5	121268	21	AC004381	Human sapiens clone D10	5.90e+01
263	17	1.5	126241	21	HS4002064	Human sapiens X22 BAC	5.90e+01
264	17	1.5	165399	21	AC004256	Human BAC clone RG033L	5.90e+01
265	17	1.5	165399	21	AC004256	Human sapiens clone D10	5.90e+01
266	17	1.5	165399	21	AC004256	Human sapiens clone D10	5.90e+01
267	17	1.5	165399	21	AC004256	Human sapiens clone D10	5.90e+01
268	17	1.5	165399	21	AC004256	Human sapiens clone D10	5.90e+01
269	17	1.5	165399	21	AC004256	Human sapiens clone D10	5.90e+01
270	17	1.5	165399	21	AC004256	Human sapiens clone D10	5.90e+01
271	17	1.5	171370	21	AC004063	Human sapiens chromosome	5.90e+01
272	17	1.5	170141	21	AC005019	Human sapiens fragile X	5.90e+01
273	17	1.5	157755	21	HP0M4P5	Human sapiens clone D10	5.90e+01
274	17	1.5	150248	13	AC004898	Human sapiens clone D10	5.90e+01
275	17	1.5	157755	21	HP0M4P5	Human sapiens clone D10	5.90e+01
276	17	1.5	173282	13	HS15344	Human DNA sequence ***	5.90e+01
277	17	1.5	173282	13	HS15344	Human DNA sequence ***	5.90e+01
278	17	1.5	200682	13	AC005019	Human sapiens clone G82	5.90e+01
279	17	1.5	286715	13	CE74593	Caenorhabditis elegans	5.90e+01
280	17	1.5	348050	13	HS1239042	Human sapiens 955 kb	5.90e+01
281	16	1.4	140720	15	HP0M4P2	Sequence 5 from patent	2.97e+02
282	16	1.4	145105	15	BP7P4	B.taurus gene fragment	2.97e+02
283	16	1.4	145105	15	BP7P4	Human vasocytic intes	2.97e+02
284	16	1.4	186200	15	HP0M4P5	Human vasocytic intes	2.97e+02
285	16	1.4	209200	15	HS3E2P7	H.sapiens CpG island D	2.97e+02
286	16	1.4	216169	15	AP0777780	Human STS Wt-5654	2.97e+02
287	16	1.4	216169	15	AP0777780	Daucus carota subsp. g	2.97e+02
288	16	1.4	216169	15	AP0777782	Daucus carota subsp. g	2.97e+02
289	16	1.4	216169	15	AP0777782	Daucus carota subsp. g	2.97e+02
290	16	1.4	216169	15	AP0777782	Daucus carota subsp. h	2.97e+02
291	16	1.4	216169	15	AP0777782	Daucus carota subsp. h	2.97e+02
292	16	1.4	216169	15	AP0777782	Daucus carota subsp. h	2.97e+02
293	16	1.4	216169	15	AP0777782	Daucus carota subsp. h	2.97e+02
294	16	1.4	216169	15	AP0777782	Daucus carota subsp. h	2.97e+02
295	16	1.4	216169	15	AP0777782	Daucus carota subsp. h	2.97e+02
296	16	1.4	216169	15	AP0777782	Daucus carota subsp. h	2.97e+02
297	16	1.4	216169	15	AP0777782	Daucus carota subsp. h	2.97e+02
298	16	1.4	30915	15	ECU674112	Equus caballus CUC DNA-C	2.97e+02
299	16	1.4	31120	15	HS153C1R	H.sapiens CPG D10	2.97e+02

US-08-007-977-9.198

300	16	1.4	323	30	ESLGVVAAH	H.sapiens rearranged I	2.97e+02
301	16	1.4	333	22	MGQI01885	Mus musculus Blozai c	2.97e+02
302	16	1.4	340	19	MGQI01885	Caesalpinia birotunda	2.97e+02
303	16	1.4	384	30	MGQI01885	Gossypium barbadense c	2.97e+02
304	16	1.4	390	19	TRUB1560	T.repens maize for ribu	2.97e+02
305	16	1.4	413	34	D43475	human STS M7-31233	2.97e+02
306	16	1.4	416	12	SPBP28E	S.pneumoniae of PBP 2B	2.97e+02
307	16	1.4	416	12	SPBP28E	S.pneumoniae of PBP 2B	2.97e+02
308	16	1.4	416	12	SPBP28E	S.pneumoniae of PBP 2B	2.97e+02
309	16	1.4	425	34	ESHRN47	H.sapiens mRNA for new	2.97e+02
310	16	1.4	434	30	ESHRN47	H.sapiens mRNA for new	2.97e+02
311	16	1.4	444	32	MSTG7VCA	Mouse (clone RAG4) T-c	2.97e+02
312	16	1.4	444	32	MSTG7VCA	Mouse (clone RAG4) T-c	2.97e+02
313	16	1.4	447	32	MSTG7VCG	Mouse (clone RAG4) T-c	2.97e+02
314	16	1.4	447	32	MSTG7VCG	Mouse (clone RAG4) T-c	2.97e+02
315	16	1.4	447	32	MSTG7VCG	Mouse (clone RAG4) T-c	2.97e+02
316	16	1.4	447	32	MSTG7VCC	Mouse (clone RAG4) T-c	2.97e+02
317	16	1.4	447	32	MSTG7VCC	Mouse (clone RAG4) T-c	2.97e+02
318	16	1.4	454	31	ESLTP2M	Mouse (clone RAG1) T-c	2.97e+02
319	16	1.4	464	12	SPBP28C	S.pneumoniae of PBP 2B	2.97e+02
320	16	1.4	491	12	SPBP28C	S.pneumoniae of PBP 2B	2.97e+02
321	16	1.4	497	12	SPBP28I	S.pneumoniae of PBP 2B	2.97e+02
322	16	1.4	508	32	MSTG7X2	Human fibroblast str 9	2.97e+02
323	16	1.4	508	32	MSTG7X2	Human fibroblast str 9	2.97e+02
324	16	1.4	517	32	MSTG7G1P	Muscle T-cell receptor	2.97e+02
325	16	1.4	517	32	MSTG7G1P	Muscle T-cell receptor	2.97e+02
326	16	1.4	542	34	G73599	SHG-58157 Human Homo	2.97e+02
327	16	1.4	542	34	G73599	SHG-58157 Human Homo	2.97e+02
328	16	1.4	562	34	G39855	EC1088 Human fibroblast	2.97e+02
329	16	1.4	571	12	HMNTG0132	Xylopiasma gentianum	2.97e+02
330	16	1.4	600	21	HMNTG01D09	Homo sapiens full leng	2.97e+02
331	16	1.4	615	34	HPD28440	Heptadactylia pseudori	2.97e+02
332	16	1.4	708	34	G320503	Heptadactylia chromom	2.97e+02
333	16	1.4	777	32	G320503	Heptadactylia chromom	2.97e+02
334	16	1.4	777	32	G320503	Heptadactylia chromom	2.97e+02
335	16	1.4	772	92	MG357395	Mus musculus class 1b	2.97e+02
336	16	1.4	777	92	MG357395	Mus musculus class 1b	2.97e+02
337	16	1.4	777	92	MG357395	Mus musculus class 1b	2.97e+02
338	16	1.4	824	30	HMNTG0132	Xylopiasma gentianum	2.97e+02
339	16	1.4	831	12	KATP7A2	Alcaligenes xyloxydant	2.97e+02
340	16	1.4	875	32	KATP7A2	Alcaligenes xyloxydant	2.97e+02
341	16	1.4	892	20	HMNTD01	Homo sapiens cytidine	2.97e+02
342	16	1.4	937	19	DCM4881	D.thalassiosira CARGES	2.97e+02
343	16	1.4	937	19	DCM4881	D.thalassiosira CARGES	2.97e+02
344	16	1.4	951	19	DINCA588	Dianthus carophyllus	2.97e+02
345	16	1.4	960	17	A36589	Sequence 3 from Patent	2.97e+02
346	16	1.4	961	19	PEPTRNS9F	Petunia transcription	2.97e+02
347	16	1.4	971	19	PEPTRNS9F	Scaphyria amara yolk	2.97e+02
348	16	1.4	971	19	PEPTRNS9F	Scaphyria amara yolk	2.97e+02
349	16	1.4	988	19	ATD30480	Arabidopsis thaliana c	2.97e+02
350	16	1.4	990	32	MSTG7G0A2	Mouse T-cell receptor	2.97e+02



351	16	1.4	936	27	RNASEP	Prospect 16	2.97e+02
352	16	1.4	1033	20	HPV6	Human papilloma virus 6	2.97e+02
353	16	1.4	1033	20	HPV6	Human papilloma virus 6	2.97e+02
354	16	1.4	1040	22	MF0214278	Mus musculus alpha-1-a	2.97e+02
355	16	1.4	1060	21	AF011524	Homo sapiens isolate T	2.97e+02
356	16	1.4	1093	15	BOVCSNA	Bovine beta-casein mRNA	2.97e+02
357	16	1.4	1112	12	SP22088	S. pneumoniae psp 28	2.97e+02
358	16	1.4	1113	12	SP22028	S. pneumoniae psp 28	2.97e+02
359	16	1.4	1250	20	ROMTICP	Human cellular fms pro	2.97e+02
360	16	1.4	1250	20	ROMTICP	Adult diarrheal toxin	2.97e+02
361	16	1.4	1386	22	AF052150	Homo sapiens T-cell	2.97e+02
362	16	1.4	1386	22	AF052150	Homo sapiens T-cell	2.97e+02
363	16	1.4	1334	16	GA0VHX14	Mus musculus clone 244	2.97e+02
364	16	1.4	1334	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
365	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
366	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
367	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
368	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
369	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
370	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
371	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
372	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
373	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
374	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
375	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
376	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
377	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
378	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
379	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
380	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
381	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
382	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
383	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
384	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
385	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
386	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
387	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
388	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
389	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
390	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
391	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
392	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
393	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
394	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
395	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
396	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
397	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
398	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
399	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02
400	16	1.4	1333	16	GA0VHX14	G.gallus avr5 gene, ex	2.97e+02

C 402	16	1-4	1480	12	SPD30069	<i>Streptococcus pneumoniae</i>	2.97e+02
C 403	16	1-4	1480	12	SPD30070	<i>Streptococcus pneumoniae</i>	2.97e+02
C 404	16	1-4	1480	12	SPD30080	<i>Streptococcus pneumoniae</i>	2.97e+02
C 405	16	1-4	1480	12	SPD20074	<i>Streptococcus pneumoniae</i>	2.97e+02
C 406	16	1-4	1480	12	SPD20070	<i>Streptococcus pneumoniae</i>	2.97e+02
C 407	16	1-4	1480	12	SPD20079	<i>Streptococcus pneumoniae</i>	2.97e+02
C 408	16	1-4	1480	12	SPD20075	<i>Streptococcus pneumoniae</i>	2.97e+02
C 409	16	1-4	1480	12	SPD20075	<i>Streptococcus pneumoniae</i>	2.97e+02
C 410	16	1-4	1488	14	SPG03281	<i>Amhyomma americanum</i>	2.97e+02
C 411	16	1-4	1505	12	SPPNR56	<i>Streptococcus pneumoniae</i>	2.97e+02
C 412	16	1-4	1505	12	SPPNR56	<i>Streptococcus pneumoniae</i>	2.97e+02
C 413	16	1-4	1512	12	SPG46874	<i>Mus musculus</i>	2.97e+02
C 414	16	1-4	1550	12	SPPER48	<i>Streptococcus pneumoniae</i>	2.97e+02
C 415	16	1-4	1551	19	STILL638	<i>Solanum tuberosum</i>	2.97e+02
C 416	16	1-4	1551	19	STILL638	<i>Solanum tuberosum</i>	2.97e+02
C 417	16	1-4	1579	31	HUNKPRA01	<i>Human cytomegalovirus</i>	2.97e+02
C 418	16	1-4	1600	37	TD127JPS	<i>Human lymphotropic vir</i>	2.97e+02
C 419	16	1-4	1601	34	TND46130	<i>Trichoplusia ni</i>	2.97e+02
C 420	16	1-4	1631	24	SPD19835	<i>Human Cytoc</i>	2.97e+02
C 421	16	1-4	1631	24	SPD19835	<i>Human Cytoc</i>	2.97e+02
C 422	16	1-4	1668	31	APF20140	<i>Hom sapiens</i>	2.97e+02
C 423	16	1-4	1707	12	SPD2LCR6M	<i>B.thuringiensis p1cr g</i>	2.97e+02
C 424	16	1-4	1723	12	MUSPR39	<i>mouse t(15;17) trans</i>	2.97e+02
C 425	16	1-4	1723	12	MUSPR39	<i>mouse t(15;17) trans</i>	2.97e+02
C 426	16	1-4	1761	31	APF38202	<i>Hom sapiens</i>	2.97e+02
C 427	16	1-4	1784	19	POTAR01A	<i>DNA encoding 3-deoxy-D-</i>	2.97e+02
C 428	16	1-4	1809	17	ET30001	<i>RNA encoding Pseudomon</i>	2.97e+02
C 429	16	1-4	1809	17	ET30001	<i>RNA encoding Pseudomon</i>	2.97e+02
C 430	16	1-4	1939	17	SPG3645	<i>Cs sequence</i>	2.97e+02
C 431	16	1-4	1842	19	VUG61396	<i>Vigna unguiculata</i>	2.97e+02
C 432	16	1-4	1852	21	SDN426	<i>H.sapiens mRNA for DAN</i>	2.97e+02
C 433	16	1-4	1852	21	SDN426	<i>H.sapiens mRNA for DAN</i>	2.97e+02
C 434	16	1-4	1910	19	DRB707	<i>Schistosom</i>	2.97e+02
C 435	16	1-4	1916	19	LEDAPSR9	<i>L.esculentum DARP synt</i>	2.97e+02
C 436	16	1-4	1932	19	ATORP02	<i>A.thaliana mRNA for un</i>	2.97e+02
C 437	16	1-4	1932	19	ATORP02	<i>A.thaliana mRNA for un</i>	2.97e+02
C 438	16	1-4	2023	17	OSCP19	<i>Human 5-lipoxygenase g</i>	2.97e+02
C 439	16	1-4	2023	17	OSCP19	<i>Human 5-lipoxygenase g</i>	2.97e+02
C 440	16	1-4	2040	19	OSCPD2	<i>O.sativa mRNA for calc</i>	2.97e+02
C 441	16	1-4	2042	21	HSENP411	<i>H.sapiens mRNA for M-p</i>	2.97e+02
C 442	16	1-4	2053	19	SGNLCB24	<i>A.cryae mRNA for glic</i>	2.97e+02
C 443	16	1-4	2107	19	SPD08290	<i>Streptococcus pneumoniae</i>	2.97e+02
C 444	16	1-4	2126	12	BBS961	<i>Bacillus sp. plasmid p</i>	2.97e+02
C 445	16	1-4	2152	22	MUSPC8	<i>Mus musculus</i>	2.97e+02
C 446	16	1-4	2152	22	MUSPC8	<i>Mus musculus</i>	2.97e+02
C 447	16	1-4	2235	22	SPD10986	<i>Mus musculus</i>	2.97e+02
C 448	16	1-4	2251	19	1YSCMK11	<i>Fest ChM1 gene for Ca</i>	2.97e+02
C 449	16	1-4	2345	19	AF072908	<i>Escolonia tabacum cal</i>	2.97e+02
C 450	16	1-4	2345	19	AF072908	<i>Human mRNA for NPI pro</i>	2.97e+02
C 451	16	1-4	2382	21	SPG11880	<i>Human mRNA for NPI pro</i>	2.97e+02
C 452	16	1-4	2393	21	HSENGMNT	<i>H.sapiens mRNA for CH</i>	2.97e+02

C 453	1.4	2300.32	MCMY1C1	Mouse normal c-myc gen	2.97e+00
C 454	1.4	2300.32	MCMY1C1	Human acidic 0.2 kDa pr	2.97e+00
C 455	1.4	2306.50	GNUS1552	Human acidic 0.2 kDa pr	2.97e+00
C 456	1.4	2436.15	GNUS1914	Glycine max galactinol	2.97e+00
C 457	1.4	2484.20	HDML05X3	Human 5-lipoxygenase M	2.97e+00
C 458	1.4	2489.20	HDML05X5	Human 5-lipoxygenase M	2.97e+00
C 459	1.4	2500.17	HDML05X6	Human 5-lipoxygenase M	2.97e+00
C 460	1.4	2500.17	HDML05X6	Human 5-lipoxygenase M	2.97e+00
C 461	1.4	2511.12	BSG49358	Caclase sequence encoding	2.97e+00
C 462	1.4	2543.32	MGC4CTPP	Caclase sequence encoding	2.97e+00
C 463	1.4	2577.17	HDML014	Mouse gamma-actin proc	2.97e+00
C 464	1.4	2577.17	HDML014	Nucleic sequence of E.	2.97e+00
C 465	1.4	2577.17	HDML014	Nucleic sequence of E.	2.97e+00
C 466	1.4	2577.17	HDML014	Nucleic sequence of E.	2.97e+00
C 467	1.4	2613.19	ASNGLA	Aspergillus oryzae DNA	2.97e+00
C 468	1.4	2621.21	HDNF1JAA	Human mRNA for NF1 N-1	2.97e+00
C 469	1.4	2627.20	HSBMT6	Human mRNA for B-myb g	2.97e+00
C 470	1.4	2632.25	HSBMT6	M. musculus B-myb RNA	2.97e+00
C 471	1.4	2648.19	SCYBR24C	S. cerevisiae chromatin	2.97e+00
C 472	1.4	2672.12	HY171219	S. cerevisiae chromatin	2.97e+00
C 473	1.4	2686.17	ES3646	Mycobacterium tubercu	2.97e+00
C 474	1.4	2713.19	YSRAN1PP	D. sequence of glucosa	2.97e+00
C 475	1.4	2713.19	YSRAN1PP	R. tabacum mRNA for ino	2.97e+00
C 476	1.4	2738.17	A4367	Sequence 1 from Paten	2.97e+00
C 477	1.4	2742.19	ES45383	Rice mRNA for vacuolar	2.97e+00
C 478	1.4	2743.22	RAREK2	Rat RRNA for vacuolar	2.97e+00
C 479	1.4	2814.32	RAREK2	Rat RRNA for vacuolar	2.97e+00
C 480	1.4	2814.32	RAREK2	Rat tyrosine protein k	2.97e+00
C 481	1.4	2866.19	DEKNA5	D. dorsalis (Nematode) m	2.97e+00
C 482	1.4	2951.12	PGNHA1	P. gingivalis nahk gene	2.97e+00
C 483	1.4	3055.19	PS14458	Plasm actinium mhaa for	2.97e+00
C 484	1.4	3127.19	SCYBR289M	S. cerevisiae chromatin	2.97e+00
C 485	1.4	3190.14	TDYBR19C	T. drosophila chromatin	2.97e+00
C 486	1.4	3190.14	TDYBR19C	T. drosophila chromatin	2.97e+00
C 487	1.4	3218.15	ES06444	glutaminase (3') region	2.97e+00
C 488	1.4	3219.15	MUSHR12NE	Mus musculus (Clone) mR	2.97e+00
C 489	1.4	3268.12	TIPOCC	Agrobacterium tumefaci	2.97e+00
C 490	1.4	3270.12	TIPOCC	Rat potassium channel-	2.97e+00
C 491	1.4	3301.22	ES04474	Rat potassium channel-	2.97e+00
C 492	1.4	3310.31	ESD84487	Human CAC chromatin p	2.97e+00
C 493	1.4	3431.21	ESD92457	Human CAC chromatin p	2.97e+00
C 494	1.4	3484.22	AF044622	Rattus norvegicus glu	2.97e+00
C 495	1.4	3505.17	AI40807X	Synthetic N7 genes N7	2.97e+00
C 496	1.4	3510.32	MUSC70XK	Mouse T cell receptor	2.97e+00
C 497	1.4	3553.12	DS8642	Gluconobacter suboxyda	2.97e+00
C 498	1.4	3576.21	ESD26814	Human ras-responsive e	2.97e+00
C 499	1.4	3599.14	DROT0012	D. melanogaster tropom	2.97e+00
C 500	1.4	3651.19	ESD7881X	S. tuberosum alpha-acti	2.97e+00
C 501	1.4	3651.19	ESD7881X	S. tuberosum alpha-acti	2.97e+00
C 502	1.4	3664.21	HSB71377	Human p58bhe (SRE) mR	2.97e+00

C	504	16	1.4	3697	22	MUSCACT	Mus musculus acyl-coen	2.97e+02
C	505	16	1.4	3704	22	RNG40001	Rattus norvegicus horn	2.97e+02
C	506	16	1.4	3710	22	RNG40002	Latrodectus polytrixus c	2.97e+02
C	507	16	1.4	3749	22	RNG40003	Mus musculus hippocamp	2.97e+02
C	508	16	1.4	3805	20	BSU37688	Human HARS1 phosphatid	2.97e+02
C	509	16	1.4	3805	15	BSU88531	Bos taurus phosphatidyl	2.97e+02
C	510	16	1.4	3884	17	A61936	Sequence 3 from Patent	2.97e+02
C	511	16	1.4	3884	17	A61936	H.sapiens mRNA for met	2.97e+02
C	512	16	1.4	3890	12	BSU0706	H.sapiens mRNA for met	2.97e+02
C	513	16	1.4	3890	12	BSU0706	White clover ribulose-	2.97e+02
C	514	16	1.4	3714	19	TRAPBXC	Flacion yeast gene for	2.97e+02
C	515	16	1.4	4118	19	ES15PSE	A.thaliana tryptophan	2.97e+02
C	516	16	1.4	4225	19	ATRRPBP	Plasmid pTAC6 from My	2.97e+02
C	517	16	1.4	4380	12	TRACOMP	Human cytochrome P-450	2.97e+02
C	518	16	1.4	4389	20	HUNEPC	Human cytic fibrosis	2.97e+02
C	519	16	1.4	4428	19	AB001348	Human alpha gene for	2.97e+02
C	520	16	1.4	4450	19	YCS0NF5	S.cerevisiae S8F5 prot	2.97e+02
C	521	16	1.4	4503	13	YCS0564	Human alpha gene for	2.97e+02
C	522	16	1.4	4601	13	OC000001	Human alpha gene for	2.97e+02
C	523	16	1.4	4866	20	HUNE740A	Human endogenous retro	2.97e+02
C	524	16	1.4	4993	12	HM7JAA15	Myobacterium tubercul	2.97e+02
C	525	16	1.4	5128	16	CRASEBP	Carassius auratus gene	2.97e+02
C	526	16	1.4	5135	12	NRACOP	A.radiobacter k12 oper	2.97e+02
C	527	16	1.4	5235	12	NRACOP	A.radiobacter k12 oper	2.97e+02
C	528	16	1.4	5448	19	DINGSTA	Dianthus caryophyllus	2.97e+02
C	529	16	1.4	5485	12	BSTR16G	Bacillus subtilis slys	2.97e+02
C	530	16	1.4	5709	22	MSH2XB	Mouse HSC class 1 Z/t1	2.97e+02
C	531	16	1.4	5709	22	MSH2XB	Mouse HSC class 1 Z/t1	2.97e+02
C	532	16	1.4	5709	22	MSH2XB	Mouse HSC class 1 Z/t1	2.97e+02
C	533	16	1.4	6089	22	MUS089D	Mus musculus class 1B	2.97e+02
C	534	16	1.4	6161	22	MUS07393	Mus musculus class 1B	2.97e+02
C	535	16	1.4	6169	22	MUS07392	Ornate adenovirus Iva2	2.97e+02
C	536	16	1.4	6314	27	MG013557	Ornate adenovirus Iva2	2.97e+02
C	537	16	1.4	6314	27	MG013557	Ornate adenovirus for	2.97e+02
C	538	16	1.4	7104	22	RNMYC	Rat c-myc oncogene and	2.97e+02
C	539	16	1.4	7235	12	RNMYC	Rat c-myc oncogene and	2.97e+02
C	540	16	1.4	7334	14	PSF0C3085	Phycomyces falciparum	2.97e+02
C	541	16	1.4	7411	12	CR9700	Myoplasma genitalium	2.97e+02
C	542	16	1.4	7400	27	SVMPDV	Swine vesicular disease	2.97e+02
C	543	16	1.4	7400	27	SVMPDV	Swine vesicular disease	2.97e+02
C	544	16	1.4	7401	27	SVMPDS	Swine vesicular disease	2.97e+02
C	545	16	1.4	7451	22	ESR09YXX	Rattus norvegicus cpmi	2.97e+02
C	546	16	1.4	7629	22	R3ID5GS	Colletotrichum legnari	2.97e+02
C	547	16	1.4	7680	19	CGP4T51	Colletotrichum legnari	2.97e+02
C	548	16	1.4	7744	14	CGM341724	Carbamiditis elegans	2.97e+02
C	549	16	1.4	7744	14	CGM341724	Carbamiditis elegans	2.97e+02
C	550	16	1.4	7941	19	SM56730	Sorghum bicolor hypothei	2.97e+02
C	551	16	1.4	8576	21	HO0246N	Human sapiens hypothei	2.97e+02
C	552	16	1.4	8694	19	CLP487XB	Picea abies f. kb. chlo	2.97e+02
C	553	16	1.4	8818	20	HM7JAA15	Myobacterium tubercul	2.97e+02
C	554	16	1.4	8918	20	HM7JAA15	Myobacterium tubercul	2.97e+02







[illegible]

C 861	15	1.3	2305	21	HUM001	Human mRNA for DBI, 21	1.35e+03
C 862	15	1.3	2140	15	RACV2P2C5	Rabbit racemase	1.35e+03
C 863	15	1.3	2163	20	HSRPEFH	Human hepatoma mRNA 20	1.35e+03
C 864	15	1.3	2404	22	MKC0D14	Mouse CD14 gene	1.35e+03
C 865	15	1.3	2405	22	HSR00160	Rus musculus neuronal	1.35e+03
C 866	15	1.3	2455	23	KL158	Human K158	1.35e+03
C 867	15	1.3	2603	14	DB8573	Plasmodium falciparum	1.35e+03
C 868	15	1.3	2643	27	ICMVBLA	Indian cavefish mosaic	1.35e+03
C 869	15	1.3	2653	26	F-ribpase gene for P5	F-ribpase gene for P5	1.35e+03
C 870	15	1.3	2655	26	HD5	Human HD5	1.35e+03
C 871	15	1.3	2708	22	RHCEPADA	Rat norvegicus mRNA	1.35e+03
C 872	15	1.3	2760	20	HUMCCPA01	Human mat c-ell carbox	1.35e+03
C 873	15	1.3	2772	14	DMPTPFSF	D melanogaster mRNA 14	1.35e+03
C 874	15	1.3	2775	14	DMPTPFSF	D melanogaster mRNA 14	1.35e+03
C 875	15	1.3	3222	22	MUSSE2632	Mouse mRNA for actin	1.35e+03
C 876	15	1.3	3398	14	AF062478	Drosophila melanogaste	1.35e+03
C 877	15	1.3	3477	14	IMG59120	Leishmania major heat	1.35e+03
C 878	15	1.3	3563	16	AA006906	Asamid for amplifying	1.35e+03
C 879	15	1.3	3566	16	AA006906	Mouse protein tyrosine	1.35e+03
C 880	15	1.3	3723	22	MUSPPTKIN	Mouse protein tyrosine	1.35e+03
C 881	15	1.3	3820	21	HUMKIAAE	Human mRNA for KIAA008	1.35e+03
C 882	15	1.3	3821	22	MUSANCI	Human mRNA for alpha N	1.35e+03
C 883	15	1.3	4003	14	DB13644	Human mRNA for KIAA001	1.35e+03
C 884	15	1.3	4603	21	DB13644	Human mRNA for KIAA001	1.35e+03
C 885	15	1.3	4635	22	RATL133A	Rat plasma proteinase	1.35e+03
C 886	15	1.3	4856	27	AF015536	JC virus type 2C, 8230	1.35e+03
C 887	15	1.3	4857	27	AF015536	JC virus type 2C, 8230	1.35e+03
C 888	15	1.3	4858	27	MVHUGA0B	M.vulva virus genome	1.35e+03
C 889	15	1.3	5111	27	U61771	Human JC virus complet	1.35e+03
C 890	15	1.3	5120	27	AF004349	JC virus variant GS/A	1.35e+03
C 891	15	1.3	5127	27	AF004350	JC virus variant Gove/1	1.35e+03
C 892	15	1.3	5173	27	HUMGLS1A	Human splein(cione 71)	1.35e+03
C 893	15	1.3	5343	20	HUMGLS1A	Human splein(cione 71)	1.35e+03
C 894	15	1.3	5657	20	HSREESTIN	H.sapiens mRNA for res	1.35e+03
C 895	15	1.3	5675	24	HSR0008	D.melanogaster polyhom	1.35e+03
C 896	15	1.3	5675	24	HSR0008	D.melanogaster polyhom	1.35e+03
C 897	15	1.3	6423	27	UKS001189	Human mRNA for ENA	1.35e+03
C 898	15	1.3	7350	21	HSR001189	Human mRNA for ENA	1.35e+03
C 899	15	1.3	7727	27	PH470C	Human papillomavirus 1	1.35e+03
C 900	15	1.3	7727	27	PH470C	Human papillomavirus 1	1.35e+03
C 901	15	1.3	10080	12	MVCSG16	Myxoma virus antigen	1.35e+03
C 902	15	1.3	10399	19	AF009516	Pentunia x hybrida tran	1.35e+03
C 903	15	1.3	10698	12	U67461	Methanococcus jamastr	1.35e+03
C 904	15	1.3	10755	14	CELI18A8	Ceniorhabditis elegans	1.35e+03
C 905	15	1.3	12300	22	KIRNXX	Rat (C. norvegicus) ei	1.35e+03
C 906	15	1.3	16300	22	KIRNXX	Rat (C. norvegicus) ei	1.35e+03
C 907	15	1.3	18006	12	D0868	E.coli genomic DNA, 6	1.35e+03
C 908	15	1.3	18072	12	HS15155	Human sapiens PHC gene	1.35e+03
C 909	15	1.3	24668	14	CEP4477	Ceniorhabditis elegans	1.35e+03
C 910	15	1.3	25689	14	CEP4477	Ceniorhabditis elegans	1.35e+03
C 911	15	1.3	26522	14	CEP72812	Ceniorhabditis elegans	1.35e+03















321 ACTTTTAACTGCGGGATGCT 339

Query Match 1.7%; Score 19; DB 22; Length 1989;  
Best Local Similarity 100.0%; Pred. No. 1.77e+00;























repeat\_region 74413..74440  
/note="7 copies of 4 mer 93 & conserved"  
repeat\_region 74932..75104  
/note="MIR repeat: matches 18..201 of consensus"  
repeat\_region 75246..75549  
/note="MIR repeat: matches 301..1 of consensus"  
repeat\_region 76433..76774  
/note="Alu repeat: matches 1..301 of consensus"  
repeat\_region 78678..79131  
/note="L1 repeat: matches 904..439 of consensus"  
repeat\_region 80131..81215  
/note="MIR repeat: matches 466..225 of consensus"  
repeat\_region 80181..80682  
/note="Alu repeat: matches 301..1 of consensus"  
repeat\_region 81131..82125  
/note="MIR repeat: matches 1927..1843 of consensus"  
repeat\_region 81852..82274  
/note="MIR repeat: matches 486..47 of consensus"  
repeat\_region 82321..82366  
/note="23 copies of 2 mer 100 & conserved"  
Note: remainder of annotations omitted.

Query Match  
Best Local Similarity 100.0%; Seed No. 1.77e+00;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 69566 AGTTCGTAAGGCACTCA 69584  
QY 329 AGTTCGTAAGGCACTCA 347

RESULT 18 AC002542 188741 bp DNA PRI 12-SEP-1997  
LOCUS AC002542 188741 bp DNA PRI 12-SEP-1997  
ACCESSION AC002542  
NID AC002542  
KEYWORDS HFG.  
SOURCE Human.  
ORGANISM Homo sapiens.  
Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 188741)  
AUTHORS Davidson S., Rohlfing T., David M. and Ahrens C.  
JOURNAL Unpublished (1997) sapiens BAC clone R014A06  
REFERENCE 2 (bases 1 to 188741)  
AUTHORS Waterston R.  
TITLE Direct Submission  
JOURNAL Unpublished (1997) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

SUBMITTED BY: Genomic Sequencing Center  
Department of Genetics  
Washington University  
St. Louis MO 63108, USA  
http://genome.wustl.edu/gsc  
mailto:seguen@wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, and the sequence was confirmed by resequencing. If a sequence was derived by sequence from more than one subclone, and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
This clone was established as part of a mapping and sequencing collaboration between the NIGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information contact John D. McPherson at jdm@wustl.edu or Eric D. Green at http://www.nigri.nih.gov/DIR/G7M/G7M/CH7 or send mail to: egreen@nigri.nih.gov

SOURCE INFORMATION:  
This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 9789K. See: Shitaya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992). Kim et al., Genomics 34:213-8 (1996). The clone is a subclone of the library. Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:  
This clone is adjacent to the right is DN056318. The actual start of this clone is at base position of R014A06. The orientation of this clone is unknown.

This clone contains STS's: sWSS377 (NID:g908836), sWSS845 (NID:g948312), and sWSS844 (NID:g916378).

Location/Qualifiers  
1..188741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

/chromosome="7"  
/clone="R014A06"  
/clone\_lib="CITB-9789K-9"  
map\_pos(2509..2530)  
complement(27181..27250)  
/rpt\_family="Li"  
27377..27368  
/note="match to EST H65143 (NID:g1023883) YU64C07.r1"  
repeat\_region complement(28075..28106)  
/rpt\_family="ALU"  
3942..4232  
/rpt\_family="ALU"  
repeat\_region complement(8246..8456)  
/rpt\_family="ALU"  
repeat\_region complement(11229..11249)  
/rpt\_family="Li"  
11804..12114  
/rpt\_family="MER"  
11958..11970  
/note="match to EST AA325774 (NID:g1978267)"  
12714..13075  
/note="match to EST AA082591 (NID:g1624650) zn3h10.r1"  
13159..13302 to EST H56178 (NID:g1190077)"  
complement(13633..13661)  
/rpt\_family="Li"  
13774..14074  
/rpt\_family="ALU"  
14759..14768  
/rpt\_family="ALU"  
14873..14945  
/rpt\_family="MER"  
14463..14925  
complement(17704..17985)  
/rpt\_family="ALU"  
complement(19061..19128)  
/rpt\_family="ALU"  
20203..20494  
/rpt\_family="ALU"  
20485..20538  
/rpt\_family="Li"  
complement(20718..21022)  
/rpt\_family="ALU"  
21573..23198  
/rpt\_family="Li"  
23539..23838  
/rpt\_family="ALU"  
24573..24665  
/rpt\_family="ALU"  
25039..25223

misc\_feature complement(2480..25620)  
/note="match to EST N7711 (NID:g1230996) YF79G08.r1"  
repeat\_region complement(27181..27250)  
/rpt\_family="Li"  
27377..27368  
/rpt\_family="ALU"  
repeat\_region complement(28075..28106)  
/rpt\_family="Li"  
29358  
/rpt\_family="ALU"  
complement(30271..30330)  
/rpt\_family="Li"  
30331..30624  
/rpt\_family="ALU"  
complement(30633..31188)  
/rpt\_family="Li"  
31845..31876  
/rpt\_family="Li"  
32454..34376  
/rpt\_family="Li"  
34692)  
complement(34701..34714)  
/rpt\_family="ALU"  
34808..35082  
/rpt\_family="ALU"  
36219..36244  
/rpt\_family="ALU"  
36584..36596  
/note="match to EST R07721 (NID:g759144) YF14B06.r1"  
36584..37002  
/note="match to EST R07091 (NID:g759014) YF13J06.r1"  
complement(37484..37714)  
/rpt\_family="ALU"  
37959..38250  
/rpt\_family="ALU"  
complement(41704..41749)  
/rpt\_family="Li"  
41819)  
/rpt\_family="Li"  
42887..43233  
/rpt\_family="MER"  
complement(43531..44074)  
/rpt\_family="ALU"  
complement(44245..44406)  
/rpt\_family="Li"  
complement(46917..47079)  
/rpt\_family="ALU"  
complement(47110..47408)  
/rpt\_family="ALU"



[illegible]

```

BASE COUNT      257 a   241 c   249 g   338 t   4 others
ORIGIN

Query Match          1.6%: Score 18; DB 19; Length 1089;
Best Local Similarity 100.0%: Prid. No. 1.07e+01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db  925 TCTGTCGTGATCACTT 942
QY  168 TCTGTCGTGATCACTT 185

RESULT 21
DEFINITION Pyruvate carboxylase, chloroplast, ribulose 1,5-bisphosphate
          carboxylase (rbcL) gene, partial cds.
ACCESSION L34779
KEYWORDS 930320
SOURCE    ribulose 1,5-bisphosphate carboxylase large subunit;
          ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit;
          rubisco large subunit.
ORIGIN    Chloroplast Pyramonas grossii (strain K-0253) DNA.
          Daugbjerg, M., Christrup, O. and Arstander, P.
          Pyramonas grossii (strain K-0253) (Pyramonaceae, Chlorophyta).
          Eukaryotes: mitochondrial eukaryotes: Viridiplantae: Chlorophyta.
          Prasinophyceae: Chlorodendraceae: Balophractaceae: Pyramoninae.
          1 (bases 1 to 1089)
          2 (bases 1 to 1089)
          3 (bases 1 to 1089)
          4 (bases 1 to 1089)
          5 (bases 1 to 1089)
          6 (bases 1 to 1089)
          7 (bases 1 to 1089)
          8 (bases 1 to 1089)
          9 (bases 1 to 1089)
          10 (bases 1 to 1089)
          11 (bases 1 to 1089)
          12 (bases 1 to 1089)
          13 (bases 1 to 1089)
          14 (bases 1 to 1089)
          15 (bases 1 to 1089)
          16 (bases 1 to 1089)
          17 (bases 1 to 1089)
          18 (bases 1 to 1089)
          19 (bases 1 to 1089)
          20 (bases 1 to 1089)
          21 (bases 1 to 1089)
          22 (bases 1 to 1089)
          23 (bases 1 to 1089)
          24 (bases 1 to 1089)
          25 (bases 1 to 1089)
          26 (bases 1 to 1089)
          27 (bases 1 to 1089)
          28 (bases 1 to 1089)
          29 (bases 1 to 1089)
          30 (bases 1 to 1089)
          31 (bases 1 to 1089)
          32 (bases 1 to 1089)
          33 (bases 1 to 1089)
          34 (bases 1 to 1089)
          35 (bases 1 to 1089)
          36 (bases 1 to 1089)
          37 (bases 1 to 1089)
          38 (bases 1 to 1089)
          39 (bases 1 to 1089)
          40 (bases 1 to 1089)
          41 (bases 1 to 1089)
          42 (bases 1 to 1089)
          43 (bases 1 to 1089)
          44 (bases 1 to 1089)
          45 (bases 1 to 1089)
          46 (bases 1 to 1089)
          47 (bases 1 to 1089)
          48 (bases 1 to 1089)
          49 (bases 1 to 1089)
          50 (bases 1 to 1089)
          51 (bases 1 to 1089)
          52 (bases 1 to 1089)
          53 (bases 1 to 1089)
          54 (bases 1 to 1089)
          55 (bases 1 to 1089)
          56 (bases 1 to 1089)
          57 (bases 1 to 1089)
          58 (bases 1 to 1089)
          59 (bases 1 to 1089)
          60 (bases 1 to 1089)
          61 (bases 1 to 1089)
          62 (bases 1 to 1089)
          63 (bases 1 to 1089)
          64 (bases 1 to 1089)
          65 (bases 1 to 1089)
          66 (bases 1 to 1089)
          67 (bases 1 to 1089)
          68 (bases 1 to 1089)
          69 (bases 1 to 1089)
          70 (bases 1 to 1089)
          71 (bases 1 to 1089)
          72 (bases 1 to 1089)
          73 (bases 1 to 1089)
          74 (bases 1 to 1089)
          75 (bases 1 to 1089)
          76 (bases 1 to 1089)
          77 (bases 1 to 1089)
          78 (bases 1 to 1089)
          79 (bases 1 to 1089)
          80 (bases 1 to 1089)
          81 (bases 1 to 1089)
          82 (bases 1 to 1089)
          83 (bases 1 to 1089)
          84 (bases 1 to 1089)
          85 (bases 1 to 1089)
          86 (bases 1 to 1089)
          87 (bases 1 to 1089)
          88 (bases 1 to 1089)
          89 (bases 1 to 1089)
          90 (bases 1 to 1089)
          91 (bases 1 to 1089)
          92 (bases 1 to 1089)
          93 (bases 1 to 1089)
          94 (bases 1 to 1089)
          95 (bases 1 to 1089)
          96 (bases 1 to 1089)
          97 (bases 1 to 1089)
          98 (bases 1 to 1089)
          99 (bases 1 to 1089)
          100 (bases 1 to 1089)
          101 (bases 1 to 1089)
          102 (bases 1 to 1089)
          103 (bases 1 to 1089)
          104 (bases 1 to 1089)
          105 (bases 1 to 1089)
          106 (bases 1 to 1089)
          107 (bases 1 to 1089)
          108 (bases 1 to 1089)
          109 (bases 1 to 1089)
          110 (bases 1 to 1089)
          111 (bases 1 to 1089)
          112 (bases 1 to 1089)
          113 (bases 1 to 1089)
          114 (bases 1 to 1089)
          115 (bases 1 to 1089)
          116 (bases 1 to 1089)
          117 (bases 1 to 1089)
          118 (bases 1 to 1089)
          119 (bases 1 to 1089)
          120 (bases 1 to 1089)
          121 (bases 1 to 1089)
          122 (bases 1 to 1089)
          123 (bases 1 to 1089)
          124 (bases 1 to 1089)
          125 (bases 1 to 1089)
          126 (bases 1 to 1089)
          127 (bases 1 to 1089)
          128 (bases 1 to 1089)
          129 (bases 1 to 1089)
          130 (bases 1 to 1089)
          131 (bases 1 to 1089)
          132 (bases 1 to 1089)
          133 (bases 1 to 1089)
          134 (bases 1 to 1089)
          135 (bases 1 to 1089)
          136 (bases 1 to 1089)
          137 (bases 1 to 1089)
          138 (bases 1 to 1089)
          139 (bases 1 to 1089)
          140 (bases 1 to 1089)
          141 (bases 1 to 1089)
          142 (bases 1 to 1089)
          143 (bases 1 to 1089)
          144 (bases 1 to 1089)
          145 (bases 1 to 1089)
          146 (bases 1 to 1089)
          147 (bases 1 to 1089)
          148 (bases 1 to 1089)
          149 (bases 1 to 1089)
          150 (bases 1 to 1089)
          151 (bases 1 to 1089)
          152 (bases 1 to 1089)
          153 (bases 1 to 1089)
          154 (bases 1 to 1089)
          155 (bases 1 to 1089)
          156 (bases 1 to 1089)
          157 (bases 1 to 1089)
          158 (bases 1 to 1089)
          159 (bases 1 to 1089)
          160 (bases 1 to 1089)
          161 (bases 1 to 1089)
          162 (bases 1 to 1089)
          163 (bases 1 to 1089)
          164 (bases 1 to 1089)
          165 (bases 1 to 1089)
          166 (bases 1 to 1089)
          167 (bases 1 to 1089)
          168 (bases 1 to 1089)
          169 (bases 1 to 1089)
          170 (bases 1 to 1089)
          171 (bases 1 to 1089)
          172 (bases 1 to 1089)
          173 (bases 1 to 1089)
          174 (bases 1 to 1089)
          175 (bases 1 to 1089)
          176 (bases 1 to 1089)
          177 (bases 1 to 1089)
          178 (bases 1 to 1089)
          179 (bases 1 to 1089)
          180 (bases 1 to 1089)
          181 (bases 1 to 1089)
          182 (bases 1 to 1089)
          183 (bases 1 to 1089)
          184 (bases 1 to 1089)
          185 (bases 1 to 1089)
          186 (bases 1 to 1089)
          187 (bases 1 to 1089)
          188 (bases 1 to 1089)
          189 (bases 1 to 1089)
          190 (bases 1 to 1089)
          191 (bases 1 to 1089)
          192 (bases 1 to 1089)
          193 (bases 1 to 1089)
          194 (bases 1 to 1089)
          195 (bases 1 to 1089)
          196 (bases 1 to 1089)
          197 (bases 1 to 1089)
          198 (bases 1 to 1089)
          199 (bases 1 to 1089)
          200 (bases 1 to 1089)
          201 (bases 1 to 1089)
          202 (bases 1 to 1089)
          203 (bases 1 to 1089)
          204 (bases 1 to 1089)
          205 (bases 1 to 1089)
          206 (bases 1 to 1089)
          207 (bases 1 to 1089)
          208 (bases 1 to 1089)
          209 (bases 1 to 1089)
          210 (bases 1 to 1089)
          211 (bases 1 to 1089)
          212 (bases 1 to 1089)
          213 (bases 1 to 1089)
          214 (bases 1 to 1089)
          215 (bases 1 to 1089)
          216 (bases 1 to 1089)
          217 (bases 1 to 1089)
          218 (bases 1 to 1089)
          219 (bases 1 to 1089)
          220 (bases 1 to 1089)
          221 (bases 1 to 1089)
          222 (bases 1 to 1089)
          223 (bases 1 to 1089)
          224 (bases 1 to 
```

```

Query Match      1.7% Score 19; DB 13; Length 195118;
Best Local Similarity 100.0%; Pred. No. 1.77e+00;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14364 GRCCTCGGCTGGCTGGG 14382
|||||
QY 143 GRCCTCGGCTGGCTGGG 161

RESULT 20
LOCUS PYCPRECLZ 1089 bp DNA PLM 08-AUG-1995
DEFINITION Pyraminomas sp. chloroplast ribulose-1,5-bisphosphate
carboxylase/oxygenase large-subunit (rbcL) gene, partial cds.
ACCESSION U00001
NID g930324
KEYWORDS ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit.
SOURCE Chloroplast Pyraminomas sp. DNA.
ORGANISM Pyraminomas sp.
Eukaryotes; Eukaryotes; Eukaryotes; Viridiplantae; Chlorophyta;
Prasinophyceae; Chlorodendraceae; Halosphaeraceae; Pyraminomas.
REFERENCE 1. (bases 1 to 1089)
AUTHORS J. Physiol. 30, 991-999 (1994)
TITLE Phylogenetic relationships of the genus P. and Aetideus P.
INFERRED FROM THE rbcL gene
JOURNAL Direct Submission
RAUBERGES W. To 1089)
JOURNAL Submitted (22-NOV-1994) Nicola Daugbjerg, Department of Phycology,
University of Copenhagen, Øster Farimagsgade 2D, 1353 Copenhagen
O, Denmark. Location/Qualifiers
FEATURES
source
1..1089
/orfname="Pyraminomas sp."
/chloroplast
/rbcL
/taxon:36998
1..1089
/gene="rbcL"
<1..>1089
/codon_start=1
/product="ribulose-1,5-bisphosphate carboxylase/oxygenase
large-subunit"-g93032s;
translation="PRAKGVRLTITPDVQAVETDLIAAPNTQVPPECCAA
VAASSTCTVTWTVTGTSLSHRTKGKTDIEPPGSENGYICTIATPIDLPESCVT
NEFTSVNTPFALRLRLLEDLISSACKTKYPGPGRIGVDLRFLATGRLPGC
ECGELIKLVANGLAKTECKYKSLFAKSRGSGRDEEVYTPQPMRDRLPTFATIRGS
LLAHIMAXMYIDGRINGIFERYAARLMSGSDHLSSCTWGKLDERRYTLGPV
DMRDAYTENDSHGISTFGDQGLPDTN".

```



[illegible]

22	RESULT	LOCUS	DEFINITION	DNA	PLN	08-AG-1995
	PTREBLH	1069 bp	Pyraminonas octopus chloroplast rribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit (rbcL) gene, partial cds.			
	9340106		1,5-bisphosphate carboxylase/oxygenase large subunit.			
	9340107		Chloroplast: Pyraminonas octopus (strain K-001) DNA.			
	9340108		Chloroplast: Pyraminonas octopus			
	9340109		Eukaryote: mitochondrial eukaryotes: Viridiplantae: Chlorophyta: Prasinophyceae: Chlorodendrakales: Halophiales: Pyraminonas.			
	9340110		1 (bases 1 to 1069)			
	9340111		Phylogeny of the genus Pyraminonas (Prasinophyceae, Chlorophyta)			
	9340112		inferred from the rbcL gene			
	9340113		J. Phycol. 30, 991-999 (1994)			

Direct Submission	1. 1089
Location/Qualifiers	for:osism="Duranimonas octopus"
source	

```

gene      /strain=K-001*
          /db_xref=taxon:36891*
          1..1089
          /gene=rbcL*
CDS       <1..>1089
          /gene=rbcL*
          /codon_start=1

```

**Tue Nov 17 08:55:24 1998**

US-08-887-977-9.190

[illegible]

```

Query Match      1.64; Score 18; DB 19; Length 1089;
Best Local Similarity 100.0%; Pred.No. 1.07e+01;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

db      925 TCTGGTGGTGCACGCTT 942
      |||||

```

FIGURE 24	DATE	BY	FOR
1	1989	BP	DNA
2	1989	BP	DNA
3	1989	BP	DNA
4	1989	BP	DNA
5	1989	BP	DNA
6	1989	BP	DNA
7	1989	BP	DNA
8	1989	BP	DNA
9	1989	BP	DNA
10	1989	BP	DNA
11	1989	BP	DNA
12	1989	BP	DNA
13	1989	BP	DNA
14	1989	BP	DNA
15	1989	BP	DNA
16	1989	BP	DNA
17	1989	BP	DNA
18	1989	BP	DNA
19	1989	BP	DNA
20	1989	BP	DNA
21	1989	BP	DNA
22	1989	BP	DNA
23	1989	BP	DNA
24	1989	BP	DNA
25	1989	BP	DNA
26	1989	BP	DNA
27	1989	BP	DNA
28	1989	BP	DNA
29	1989	BP	DNA
30	1989	BP	DNA
31	1989	BP	DNA
32	1989	BP	DNA
33	1989	BP	DNA
34	1989	BP	DNA
35	1989	BP	DNA
36	1989	BP	DNA
37	1989	BP	DNA
38	1989	BP	DNA
39	1989	BP	DNA
40	1989	BP	DNA
41	1989	BP	DNA
42	1989	BP	DNA
43	1989	BP	DNA
44	1989	BP	DNA
45	1989	BP	DNA
46	1989	BP	DNA
47	1989	BP	DNA
48	1989	BP	DNA
49	1989	BP	DNA
50	1989	BP	DNA
51	1989	BP	DNA
52	1989	BP	DNA
53	1989	BP	DNA
54	1989	BP	DNA
55	1989	BP	DNA
56	1989	BP	DNA
57	1989	BP	DNA
58	1989	BP	DNA
59	1989	BP	DNA
60	1989	BP	DNA
61	1989	BP	DNA
62	1989	BP	DNA
63	1989	BP	DNA
64	1989	BP	DNA
65	1989	BP	DNA
66	1989	BP	DNA
67	1989	BP	DNA
68	1989	BP	DNA
69	1989	BP	DNA
70	1989	BP	DNA
71	1989	BP	DNA
72	1989	BP	DNA
73	1989	BP	DNA
74	1989	BP	DNA
75	1989	BP	DNA
76	1989	BP	DNA
77	1989	BP	DNA
78	1989	BP	DNA
79	1989	BP	DNA
80	1989	BP	DNA
81	1989	BP	DNA
82	1989	BP	DNA
83	1989	BP	DNA
84	1989	BP	DNA
85	1989	BP	DNA
86	1989	BP	DNA
87	1989	BP	DNA
88	1989	BP	DNA
89	1989	BP	DNA
90	1989	BP	DNA
91	1989	BP	DNA
92	1989	BP	DNA
93	1989	BP	DNA
94	1989	BP	DNA
95	1989	BP	DNA
96	1989	BP	DNA
97	1989	BP	DNA
98	1989	BP	DNA
99	1989	BP	DNA
100	1989	BP	DNA

**AUTHORS** Daugbjerg, N.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (22-NOV-1994) Niels Daugbjerg, Department of Physiology,  
University of Copenhagen, Øster Farimagsgade 2D, 1353 Copenhagen  
0, Denmark  
**FEATURES** Location/Qualifiers  
**source** 1 1089

## FEATURES

**Tue Nov 17 08:55:24 1998**

US-08-887-977-9.190

[illegible]

BASE COUNT 273 a 212 c 442 g 362 t  
ORIGIN  
Query Match  
Best Local similarity 100.0: Pred. NO. 1.07e+01:  
1.68: Score 18: DB 19: Length 1089:  
LEETRAMVVDYRNRHGFVZFKALRMGGDHLHGCVVGLKEDEVTGLG  
DLKRAVLENDKSGVYFDVCGCGLGSH+

DB	925	TCTGTGGTGATCACTT	942
QY	168	TCTGTGGTGATCACTT	185

RESULT	25	AB093169	1302 bp	mRNA	RCD	09-DEC-1997
DEFINITION		Mus musculus mRNA for G protein-coupled receptor Kv4.1l, complete cds.				
ACCESSION		U095149				
NID		Q095149				
KEYWORDS		G protein-coupled receptor Kv4.1l.				
SOURCE		Mus musculus cDNA to mRNA.				
ORGANISM		Mus musculus				
		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
		Rodentia; Sciurognathi; Muridae; Murinae; Mus.				

TITLE	Mouse G protein-coupled receptor K411
JOURNAL	Published Only in Database (1997) In press
REFERENCE	2 (bases 1 to 1302)
AUTHORS	Tanagihara, S., Komura, E. and Yamaguchi, T.
TITLE	Direct sublimation
JOURNAL	submitted (01-DEC-1997) to the DDBJ/EBI/GenBank databases.

**TITLE**  
**JOURNAL**

**EMBL/GenBank databases.**















Page 83

D	b	13700	TTCCTCGTTTATGAAG	13717
Q	y	184	TTCGCTTTTATGAAG	201
<hr/>				
RESULT	35	067522	19436 bp	DNA
LOCUS		Methanococcus jannaschii section 64 of 150 of the complete genome		BCT
DEFINITION		067522 L77117		
ACCESSION		9426545		
KEYWORDS				
SOURCE		Methanococcus jannaschii.		
ORGANISM		Methanococcus jannaschii.		
REFERENCE		1 (bases 1 to 19436)		
AUTHORS		Rutt,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D., Sutton,G.G., Blake,J.A., Visselard,L.M., Clayton,R.A., Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Neinstock,K.G., Herrick,M.K., Glodek,A., Scott,J.D., Geoghagan,N.S., Weidman,J.F., Fuhrman,J.L., Nguyen,D.T., Uitterback,T., Kelley,J.H., McManis,P.E., Klenk,H.P., Borkovskiy,M., Klenk,H.P., Fraser,C.M., Roberts,K.M., Kaine,S.B., Borodovsky,M., Klenk,H.P., Fraser,C.M., Smith,B.O., Weiss,C.R. and Venter,J.C.		
TITLE		Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.		
JOURNAL		Science 273 (5278), 1508-1073 (1996)		
MEDLINE		96337999		
REFERENCES		2 (bases 1 to 19436)		
AUTHORS		Rutt,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D., Gokayne,J.B., Karlagaye,A.R., Dougherty,B., Tomb,J., Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Neinstock,K.G., Herrick,M.K., Glodek,A., Scott,J.D., Geoghagan,N.S., Weidman,J.F., Fuhrman,J.L., Nguyen,D.T., Uitterback,T., Kelley,J.H., McManis,P.E., Klenk,H.P., Borkovskiy,M., Klenk,H.P., Fraser,C.M., Roberts,K.M., Kaine,S.B., Borodovsky,M., Klenk,H.P., Fraser,C.M., Smith,B.O., Weiss,C.R. and Venter,J.C.		
TITLE		Direct Submission--1996 The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA		
FEATURES		Location/Qualifiers		
source		1..19436		
		/db_xref=taxon:2190*		'Methanococcus jannaschii'
gene		/db_xref=taxon:2190*		
		/complement(124..1770)		
CDS		/gene=MJ0765*		
		/complement(124..1770)		
		/note=similar to GB:181975 SP:001770 PID:49286 percent		























[illegible]











US-08-087-977-9.19e

**Tue Nov 17 08:55:24 1998**

US-08-887-977-9, 199

```

45
RESULT      H52877L14 131187 bp      DNA      H7C      17-JUL-1998
LOCUS
DEFINITION  Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone
            H52877, H7C5 Phase 1.
ACCESSION  G55125
RID        G3334691
SOURCE     H7C: H7C5, PHASE1.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  Submitted (17-JUL-1998) Wellcome Trust Genome Campus, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            genbank@wtp.sanger.ac.uk
            IMPORTANT: This sequence is unfinished and does not necessarily
            represent the correct sequence. Work on the sequence is in progress
            and the release of this data is based on the understanding that the
            contaminated with foreign work continues. The sequence may be
            phase etc. Order of segments is not known; 800 n's separate
            segments. Unfinished: GJ2877L14 Contig.ID: G2145 acc=G55125
            Length: 10229 bp. Unfinished: GJ2877L14 Contig.ID: G2193 acc=G55125
            Length: 97146 bp. Unfinished: GJ2877L14 Contig.ID: G2886 acc=G55135
            ***
            *** WARNING: Phase 1 High Throughput Genome Sequence ***
            * This sequence is unfinished. When sequence is complete,
            * the sequence data presented in this record will be replaced
            * by a single finished sequence with the same accession number.
FEATURES
    source
        1..131187
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="2p21.1"
BASE COUNT  37993 a 20501 c 19457 g 33636 t 1600 others
ORIGIN
Query Match      1.5%: Score 18, DB 131. Length 131187;
Best Local Similarity 100.0%: Pred. No. 1,07e+01;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 77579 TCTTGAGGACCTGTGGCT 77596
QY 762 TCTTGAGGACCTGTGGCT 979

```

**Tue Nov 17 08:55:24 1998**

US-08-887-977-9.rge

[illegible][illegible]



```

The Row 17 08:55:24 1998
US-08-887-977-9.sye

repeat_region complement(20212, .20549)
repeat_region /note="LIP13 repeat: matches 300, .570 of consensus"
repeat_region complement(20050, .20846)
repeat_region /note="LIP13 repeat: matches 300, .1 of consensus"
repeat_region complement(21104, .21286)
repeat_region /note="LIP13 repeat: matches 300, .762 of consensus"
repeat_region complement(21104, .21286)
repeat_region /note="LIP13 repeat: matches 300, .762 of consensus"
repeat_region complement(21193, .21582)
repeat_region /note="LIP13 repeat: matches 300, .128 of consensus"
repeat_region complement(21431, .21772)
repeat_region /note="LIP13 repeat: matches 300, .570 of consensus"
repeat_region complement(21773, .22069)
repeat_region /note="LIP13 repeat: matches 300, .1 of consensus"
repeat_region complement(22170, .22823)
repeat_region /note="LIP13 repeat: matches 300, .362 of consensus"
repeat_region complement(22177, .22509)
repeat_region /note="LIP13 repeat: matches 300, .762 of consensus"
repeat_region complement(22516, .22805)
repeat_region /note="LIP13 repeat: matches 300, .128 of consensus"
repeat_region complement(22823, .23112)
repeat_region /note="LIP13 repeat: matches 300, .128 of consensus"
repeat_region /note="AluX repeat: matches 300, .7 of consensus"
repeat_region 23501, .23544
repeat_region /note="MIR repeat: matches 56, .98 of consensus"
repeat_region 41131, .41235
repeat_region /note="MIR repeat: matches 6, .201 of consensus"
repeat_region complement(24144, .24436)
repeat_region /note="AluX repeat: matches 300, .7 of consensus"
repeat_region 24724, .24767
repeat_region /note="LIP13 repeat: matches 56, .98 of consensus"
repeat_region complement(25522, .25551)
repeat_region /note="LIP13 repeat: matches 300, .762 of consensus"
repeat_region /note="LIP13 repeat: matches 5021, .4889 of consensus"
repeat_region 25356, .25546
repeat_region /note="MIR repeat: matches 6, .201 of consensus"
repeat_region complement(25546, .25551)
repeat_region /note="AluX repeat: matches 300, .762 of consensus"
repeat_region complement(26444, .26576)
repeat_region /note="LIP13 repeat: matches 300, .128 of consensus"
repeat_region complement(27192, .27534)
repeat_region /note="LIP13 repeat: matches 300, .128 of consensus"
repeat_region complement(27178, .27672)
repeat_region /note="AluX repeat: matches 345, .1 of consensus"
repeat_region complement(27763, .28023)
repeat_region /note="LIP13 repeat: matches 300, .128 of consensus"
repeat_region /note="LIP13 repeat: matches 1964, .1709 of consensus"
repeat_region /note="MIR repeat: matches 345, .1 of consensus"
repeat_region complement(28886, .29246)
repeat_region /note="MIR repeat: matches 6, .201 of consensus"
...
Query Match
Score 1.69; Score 18; DB 21; Length 173513;

```

```

Run Row 17 00:55:24 1998

US-08-887-977-9.9pg

/4b.tref=:taxon:9606"
/Chromosome="21"
/Clone="211A9"
BASE COUNT 44710 a 39974 c 40165 g 48034 t 7280 others
GALVIN

Query Match 1.6% Score 18; DB 13; Length 180163;
Best Local Similarity 100.0%; Pred. No. 1.07e+01;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 51664 TGCTGCTGCTGCTGCTG 51681
|||||
QY 749 TGCTGCTGCTGCTGCTG 765

RESULT 49
LOCUS AC004479 190000 bp DNA HTG 03-JUL-1998
DEFINITION *** SEQUENCING IN PROGRESS *** Homo sapiens chromosome 4, BAC clone
COSMID: HTCS Phase 1, 48 unordered pieces.
ACCESSION G1287659
XREF NID
KEYWORDS HTG; HTGS; PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Chordata; Chordata; Vertebrata; Mammalia; Eutheria.
Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 190000)
Stone, N.E., Schatz, J.J., Cox, D.R. and Myers, R.M.
Direct Subsequencing
JOURNAL
2 (bases 1 to 190000)
Stone, N.E., Schatz, J.J., Cox, D.R. and Myers, R.M.
Direct Subsequencing
JOURNAL
Submitted (27-MAR-1998) Department of Genetics, Stanford Human
Genome Center, 85 California Avenue, Palo Alto, CA 94304, USA
*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
This sequence is unfinished. It consists of 48 contigs for
which the exact lengths are unknown. The exact lengths of the gaps
between the contigs are also unknown; these gaps are presented
as runs of N as a convenience only. When sequencing is complete,
the sequence data presented in this record will be replaced
by a single 19048 bp contig of 1048 bp in length.
1 1049 1195: gap of unknown length
2 1196 2336: contig of 1121 bp in length
3 2427 3463: gap of unknown length
4 3464 3547: gap of unknown length
5 3547 3693: gap of unknown length
COMMENT

```















C 349	15	1.3	2559	20	710815	Kaposi's sarcoma asso	8.37e+01
C 350	14	1.3	2560	21	710816	Metastatic sarcoma	8.37e+01
C 351	14	1.3	2561	22	710817	Insulin-stimulated pr	3.47e+02
C 352	14	1.3	2562	23	710818	T. niteum GAPDH	3.47e+02
C 353	14	1.3	2563	24	710819	Myobacterial abcp o	3.47e+02
C 354	14	1.3	2564	25	710820	Human caprin-2	3.47e+02
C 355	14	1.3	2565	26	710821	Human caprin-2	3.47e+02
C 356	14	1.3	2566	27	710822	Sequence encoding esp	3.47e+02
C 357	14	1.3	2567	28	710823	Staphylococcus aureus	3.47e+02
C 358	14	1.3	2568	29	710824	Drosophila frizzled g	3.47e+02
C 359	14	1.3	2569	30	710825	MyoD myofibrinogen	3.47e+02
C 360	14	1.3	2570	31	710826	MyoD myofibrinogen	3.47e+02
C 361	14	1.3	2571	32	710827	Nucleotide sequence o	3.47e+02
C 362	14	1.3	2572	33	710828	5' region of human ce	3.47e+02
C 363	14	1.3	2573	34	710829	S. aureus topoisomeras	3.47e+02
C 364	14	1.3	2574	35	710830	Glutathione transfera	3.47e+02
C 365	14	1.3	2575	36	710831	Glutathione transfera	3.47e+02
C 366	14	1.3	2576	37	710832	Sequence comprising a	3.47e+02
C 367	14	1.3	2577	38	710833	Rat cryptid 2 gene.	3.47e+02
C 368	14	1.3	2578	39	710834	Marine FRC X1 isozyme	3.47e+02
C 369	14	1.3	2579	40	710835	Partial cDNA, residue	3.47e+02
C 370	14	1.3	2580	41	710836	Partial cDNA molecule	3.47e+02
C 371	14	1.3	2581	42	710837	Sequence encoding pro	3.47e+02
C 372	14	1.3	2582	43	710838	Rat cryptid 1 gene.	3.47e+02
C 373	14	1.3	2583	44	710839	Sequence encoding pro	3.47e+02
C 374	14	1.3	2584	45	710840	Rat cryptid 1 gene.	3.47e+02
C 375	14	1.3	2585	46	710841	MRV-1 clone LTR/G12	3.47e+02
C 376	14	1.3	2586	47	710842	Fragment pHS 53 hmo1	3.47e+02
C 377	14	1.3	2587	48	710843	cDNA encoding human F	3.47e+02
C 378	14	1.3	2588	49	710844	Rat cryptid 3 gene.	3.47e+02
C 379	14	1.3	2589	50	710845	Sequence encoding cel	3.47e+02
C 380	14	1.3	2590	51	710846	Sequence of Trypanoso	3.47e+02
C 381	14	1.3	2591	52	710847	Pistum sativum L PAL	3.47e+02
C 382	14	1.3	2592	53	710848	Enocodes T3 RNA polye	3.47e+02
C 383	14	1.3	2593	54	710849	Enocodes T3 RNA polye	3.47e+02
C 384	14	1.3	2594	55	710850	Enocodes T3 RNA polye	3.47e+02
C 385	14	1.3	2595	56	710851	Enocodes T3 RNA polye	3.47e+02
C 386	14	1.3	2596	57	710852	Enocodes T3 RNA polye	3.47e+02
C 387	14	1.3	2597	58	710853	Enocodes T3 RNA polye	3.47e+02
C 388	14	1.3	2598	59	710854	Enocodes T3 RNA polye	3.47e+02
C 389	14	1.3	2599	60	710855	Enocodes T3 RNA polye	3.47e+02
C 390	14	1.3	2600	61	710856	Enocodes T3 RNA polye	3.47e+02
C 391	14	1.3	2601	62	710857	Enocodes T3 RNA polye	3.47e+02
C 392	14	1.3	2602	63	710858	Enocodes T3 RNA polye	3.47e+02
C 393	14	1.3	2603	64	710859	Enocodes T3 RNA polye	3.47e+02
C 394	14	1.3	2604	65	710860	Enocodes T3 RNA polye	3.47e+02
C 395	14	1.3	2605	66	710861	Enocodes T3 RNA polye	3.47e+02
C 396	14	1.3	2606	67	710862	Enocodes T3 RNA polye	3.47e+02
C 397	14	1.3	2607	68	710863	Enocodes T3 RNA polye	3.47e+02
C 398	14	1.3	2608	69	710864	Enocodes T3 RNA polye	3.47e+02
C 399	14	1.3	2609	70	710865	Enocodes T3 RNA polye	3.47e+02

C 451	14	1.3	3738	3	Q14312	Bacillus thuringiensis	3.47e+02
C 452	14	1.3	3739	4	Q14313	Marine tyrosine kinase	3.47e+02
C 453	14	1.3	3740	5	Q14314	c-abl gene.	3.47e+02
C 454	14	1.3	3741	6	Q14315	Human haematopoietic	3.47e+02
C 455	14	1.3	3742	7	Q14316	Amplified precursor	3.47e+02
C 456	14	1.3	3743	8	Q14317	Genomic DNA clone	3.47e+02
C 457	14	1.3	3744	9	Q14318	Bacillus thuringiensis	3.47e+02
C 458	14	1.3	3745	10	Q14319	BT toxin 17b.	3.47e+02
C 459	14	1.3	3746	11	Q14320	BT toxin 17b.	3.47e+02
C 460	14	1.3	3747	12	Q14321	BT toxin 17b.	3.47e+02
C 461	14	1.3	3748	13	Q14322	BT toxin 17b.	3.47e+02
C 462	14	1.3	3749	14	Q14323	Bacillus thuringiensis	3.47e+02
C 463	14	1.3	3750	15	Q14324	17b toxin coding sequ	3.47e+02
C 464	14	1.3	3751	16	Q14325	Bacillus thuringiensis	3.47e+02
C 465	14	1.3	3752	17	Q14326	Bacillus thuringiensis	3.47e+02
C 466	14	1.3	3753	18	Q14327	Gene encoding C-termi	3.47e+02
C 467	14	1.3	3754	19	Q14328	Enocodes equine C-termi	3.47e+02
C 468	14	1.3	3755	20	Q14329	Sequence of plasmid p	3.47e+02
C 469	14	1.3	3756	21	Q14330	Human WR-1720	3.47e+02
C 470	14	1.3	3757	22	Q14331	Human WR-1720	3.47e+02
C 471	14	1.3	3758	23	Q14332	Human WR-1720	3.47e+02
C 472	14	1.3	3759	24	Q14333	Human WR-1720	3.47e+02
C 473	14	1.3	3760	25	Q14334	Human WR-1720	3.47e+02
C 474	14	1.3	3761	26	Q14335	Human WR-1720	3.47e+02
C 475	14	1.3	3762	27	Q14336	Human WR-1720	3.47e+02
C 476	14	1.3	3763	28	Q14337	Human WR-1720	3.47e+02
C 477	14	1.3	3764	29	Q14338	Human WR-1720	3.47e+02
C 478	14	1.3	3765	30	Q14339	Human WR-1720	3.47e+02
C 479	14	1.3	3766	31	Q14340	Human WR-1720	3.47e+02
C 480	14	1.3	3767	32	Q14341	Human WR-1720	3.47e+02
C 481	14	1.3	3768	33	Q14342	Human WR-1720	3.47e+02
C 482	14	1.3	3769	34	Q14343	Human WR-1720	3.47e+02
C 483	14	1.3	3770	35	Q14344	Human WR-1720	3.47e+02
C 484	14	1.3	3771	36	Q14345	Human WR-1720	3.47e+02
C 485	14	1.3	3772	37	Q14346	Human WR-1720	3.47e+02
C 486	14	1.3	3773	38	Q14347	Human WR-1720	3.47e+02
C 487	14	1.3	3774	39	Q14348	Human WR-1720	3.47e+02
C 488	14	1.3	3775	40	Q14349	Human WR-1720	3.47e+02
C 489	14	1.3	3776	41	Q14350	Human WR-1720	3.47e+02
C 490	14	1.3	3777	42	Q14351	Human WR-1720	3.47e+02
C 491	14	1.3	3778	43	Q14352	Human WR-1720	3.47e+02
C 492	14	1.3	3779	44	Q14353	Human WR-1720	3.47e+02
C 493	14	1.3	3780	45	Q14354	Human WR-1720	3.47e+02
C 494	14	1.3	3781	46	Q14355	Human WR-1720	3.47e+02
C 495	14	1.3	3782	47	Q14356	Human WR-1720	3.47e+02
C 496	14	1.3	3783	48	Q14357	Human WR-1720	3.47e+02
C 497	14	1.3	3784	49	Q14358	Human WR-1720	3.47e+02
C 498	14	1.3	3785	50	Q14359	Human WR-1720	3.47e+02
C 499	14	1.3	3786	51	Q14360	Human WR-1720	3.47e+02
C 500	14	1.3	3787	52	Q14361	Human WR-1720	3.47e+02
C 501	14	1.3	3788	53	Q14362	Human WR-1720	3.47e+02























QY 935 AGAGTTTCAAGACTAC 951

RESULT 10

ID T80118 standard; cDNA: 1442 BP.

AC T80118-1998 (first entry)

DE cDNA for human CCR5.

DE Human Cys-Cys chemokine receptor 5; CCR5.

DE Human immunodeficiency virus type 1; HIV-1; HIV-2.

KW Human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2.

KW Inactive human Cys-Cys chemokine receptor 5; CCR5.

KW predisposition; resistance; diagnosis; treatment; prevention;

KW cancer; idiopathic pulmonary fibrosis; parosmia; viral infection;

KW cancer; atherosclerosis; autoimmune disorder; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 240..1398

FT /tag= a

WT W09712019-A2.

PD 04-SEP-1997. B00023

PR 04-SEP-1997. B00023

PR 06-AUG-1996; EP-870102.

PR 01-MAR-1996; EP-870021.

PA (EURO-) EUROSCREEN SA.

PI Libert F, Parmentier M, Samson M, Vassart G;

DR P-PSDB: 42782/44.

DR P-PSDB: 790117.

PT Active and inactive forms of human CC chemokine receptor CCR-5 -

PT useful to diagnose, prevent and/or treat inflammatory disorders,

PT autoimmune disease and viral infection

CC The present sequence encodes human CC (Cys-Cys) chemokine receptor

CC chemokine receptor 5 (CCR5), which lacks the last 3 transmembrane

CC regions and the regions involved in G protein-coupling. CCR5 or

CC its cDNA can be used to diagnose, treat and/or prevent inflammatory

CC diseases, such as rheumatoid arthritis, fibrosis, atherosclerosis,

CC idiopathic pulmonary fibrosis and parosmia. It is also a receptor of

CC especially human immunodeficiency virus type 1 or type 2 (HIV-1 or

CC HIV-2) infection. Cancer, atherosclerosis and autoimmune disorders.

CC Subjects that express the inactive receptor have a predisposition,

CC cancer, atherosclerosis and autoimmune disorders.

CC Sequence 1442 BP; 314 A; 343 C; 422 T;

SQ

Query Match 1.5%; Score 17; DB 37; Length 1442;

Best Local Similarity 100.0%; Pred. No. 3.85e+00;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1112 agagtttcaagactac 1128

QY 935 AGAGTTTCAAGACTAC 951

RESULT 11

ID T80117 standard; cDNA: 1477 BP.

AC T80117-1998 (first entry)

DE cDNA for human CCR5.

DE Human Cys-Cys chemokine receptor 5; CCR5.

DE Human immunodeficiency virus type 1; HIV-1; HIV-2.

KW Human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2.

KW Inactive human Cys-Cys chemokine receptor 5; CCR5.

KW predisposition; resistance; diagnosis; treatment; prevention;

KW cancer; idiopathic pulmonary fibrosis; parosmia; viral infection;

KW cancer; atherosclerosis; autoimmune disorder; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 240..1398

FT /tag= a

WT W09712019-A2.

PD 04-SEP-1997. B00023

PR 06-AUG-1996; EP-870102.

PR 01-MAR-1996; EP-870021.

PA (EURO-) EUROSCREEN SA.

PI Libert F, Parmentier M, Samson M, Vassart G;

DR P-PSDB: 42782/44.

DR P-PSDB: 790117.

PT Active and inactive forms of human CC chemokine receptor CCR-5 -

PT useful to diagnose, prevent and/or treat inflammatory disorders,

PT autoimmune disease and viral infection

CC The present sequence encodes human CC (Cys-Cys) chemokine receptor

CC chemokine receptor 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES

CC (MIP-2). MIP-1, but not by monocyte chemoattractant protein 1 (MCP-1),

CC MIP-2, MCP-1, is involved in the pathogenesis of AIDS. It is also a

CC alpha (GRO alpha) chemokine. Active CCR-5 is also a receptor of

CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).

CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent

CC inflammatory diseases, such as rheumatoid arthritis, fibrosis and

CC glomerulonephritis. It is also a receptor of especially HIV-1 or HIV-2 infection,

CC parosmia, viral infections, especially HIV-1 or HIV-2 infection,

CC cancer, atherosclerosis and autoimmune disorders.

CC Sequence 1477 BP; 374 A; 320 G; 431 T;

SQ

Query Match 1.5%; Score 17; DB 37; Length 1477;

Best Local Similarity 100.0%; Pred. No. 3.85e+00;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1144 agagtttcaagactac 1160

QY 935 AGAGTTTCAAGACTAC 951

RESULT 12

ID T80107 standard; cDNA: 1526 BP.

AC T80107-1997 (first entry)

DE Human NF-1 gene; cDNA: KP361-2.

DE Human NF-1 gene; cDNA: KP361-2.

KW Alzheimer's disease; NF-1; human; mouse; KP361 gene; antibody; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 361..1526

FT /tag= a

WT J0213495-A.

PD 10-JUN-1996; JP-345659.

PR 07-DEC-1995; JP-345659.

PA (SDSU) SUMITOMO SEIYAKU KK.

PI NF-1; 97-46507/345659; 181514.

DR P-PSDB: 42782/44.

DR P-PSDB: 790117.

PT Active and inactive forms of the human NF-1 gene (see T80099 for

PT wild type sequence). This sequence, T80098, and T80100 all represent cDNA

PT clones of the human NF-1 gene. The NF-1 gene is a tumor suppressor

PT gene. The NF-1 gene is useful for the diagnosis of Alzheimer's

PT disease.

CC Example 3; Page 15-16; 23pp; Japanese.

CC T80106 and T80107 represent clones of the human NF-1 gene (see T80099 for

CC wild type sequence). This sequence, T80098, and T80100 all represent cDNA

CC clones of the human NF-1 gene. The NF-1 gene is a tumor suppressor

CC gene. The NF-1 gene is useful for the diagnosis of Alzheimer's

CC disease.

CC Sequence 1526 BP; 436 A; 269 C; 315 G; 506 T;

SQ

Query Match 1.5%; Score 17; DB 34; Length 1526;

Best Local Similarity 100.0%; Pred. No. 3.85e+00;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1015 tggcttctatataaa 1031

Cp 205 TGGCTTCTATATAAA 189

RESULT 13

ID T80952 standard; cDNA: 1557 BP.

AC T80952-1998 (first entry)

DE Human chemokine receptor MMR-CCR cDNA.

DE Human chemokine receptor MMR-CCR cDNA.

KW human; monocyte; macrophage; chemokine; chemokine receptor; MMR-CCR;

KW Inflammation; proliferative disease; cardiovascular disease;

KW tumor; rheumatoid arthritis; alveolitis; atherosclerosis;

KW chronic granulomatous disease; asthma; myasthenia gravis;

KW septic shock; Chediak-Higashi syndrome; therapy; diagnosis; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 361..1557

FT /tag= a

WT W0971225-A2.

PD 06-NOV-1997.

PR 25-APR-1997; U06993.

RESULT 14

ID T80102 standard; cDNA: 1641 BP.

AC T80102-1997 (first entry)

DE Human NF-1 gene open reading frame.

DE Human NF-1 gene open reading frame.

KW Alzheimer's disease; NF-1; human; mouse; KP361 gene; antibody; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 361..1641

FT /tag= a

WT J0213495-A.

PD 10-JUN-1996; JP-345659.

PR 07-DEC-1995; JP-345659.

PA (SDSU) SUMITOMO SEIYAKU KK.

PI NF-1; 97-46507/345659; 181514.

DR P-PSDB: 42782/44.

DR P-PSDB: 790117.

PT Active and inactive forms of the human NF-1 gene (see T80099 for

PT wild type sequence). This sequence, T80098, and T80100 all represent cDNA

PT clones of the human NF-1 gene. The NF-1 gene is a tumor suppressor

PT gene. The NF-1 gene is useful for the diagnosis of Alzheimer's

PT disease.

CC Example 3; Page 15-16; 23pp; Japanese.

CC T80106 and T80107 represent clones of the human NF-1 gene (see T80099 for

CC wild type sequence). This sequence, T80098, and T80100 all represent cDNA

CC clones of the human NF-1 gene. The NF-1 gene is a tumor suppressor

CC gene. The NF-1 gene is useful for the diagnosis of Alzheimer's

CC disease.

CC Sequence 1557 BP; 419 A; 348 C; 327 G; 463 T;

SQ

Query Match 1.5%; Score 17; DB 39; Length 1557;

Best Local Similarity 100.0%; Pred. No. 3.85e+00;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 880 agagtttcaagactac 896

QY 935 AGAGTTTCAAGACTAC 951











US-08-887-977-9.100

KW	resistant pneumococci"
OS	Penicillin resistant pneumococcus.
FH	Location/Qualifiers
Key	1..777
FT	/note= "s
FT	/note= "does not contain start or stop codon"
FT	220..237
FT	/tag= b
FT	/note= "region preferred in probe for penicillin-
PN	708098699-A.
PT	16-APR-1996.
PP	03-OCT-1994; 238305.
PR	03-OCT-1994; JP-238305.
PR	03-OCT-1994; JP-238305.
WP:	WP: 36-245875/25
DR	P-RSD: R98901.
DR	Penicillin-resistant pneumococcus gene fragment - used to detect
DR	penicillin-resistant strains of pneumococcus
PC	Class I penicillin-resistant strains of pneumococcus
CC	This fragment is a fragment of a penicillin-resistant pneumococcus gene. The
CC	fragment is useful as a primer/probe for the detection of pneumococcal
CC	existing penicillin-resistance. A sequence of 18 nucleotides (220
CC	nucleotides) was determined from the DNA of strain 98901.
CC	The present within the primer/probe derived from the above fragment. The
CC	probe allows exact and rapid diagnosis of penicillin resistance.
SQ	Sequence 777 BP: 204 A: 163 C: 183 G: 227 T;
	Query Match 1..49; Score 16; DB 21; Length 777;
	Best local Similarity 100.0%; Pred.No. 1,86e+01;
Matches	16; Conservative 0; Mismatches 0; Indels 0; Gaps
Dbl	708 atgcacaccagctgtt 723
Cp	703 cttgacccaggcttt 688
	RESULT 26
ID	T61423 standard; cdna: 892 BP.
AC	T61423;
DT	13-NOV-1997 (first entry)
DT	13-NOV-1997 (last entry)
KN	Cytidine deaminase; cytosine analogue; cytosine arabinoside;
KW	cancer; gene therapy; retrovirus; vector; pMG-CD; pOR-2;
KW	pGR-CB; bone marrow; selectable marker; human; ss.
FT	Xeno sapiens.
FT	cds 118..558
FT	Location/Qualifiers
FT	"tag= a
PN	N05705254-A1.
PT	19-JUL-1996.
PP	19-JUL-1996; CM494.

118-08-887-977-9-unc

[illegible]







CC preadherent properties for circulating T cells and monocytes. A secreted  
CC or soluble form, consisting of the chemokine domain and the stalk region,  
CC is able to inhibit this chemotactic activity. This suggests that the  
CC membrane bound form of CX3C chemokine may be a potent regulator of  
CC circulating immune cells. The soluble form may be used in the treatment  
CC of diseases, especially in conditions of compromised immune response.  
CC CX3C chemokines may also be involved in the pathogenesis of such CNS  
CC inflammatory disorders as multiple sclerosis, and other pathologies  
CC disorders associated with CX3C chemokine misregulation. They can also be  
CC used in the treatment of conditions associated with abnormal  
CC proliferation, e.g. cancerous conditions, or degenerative conditions.  
CC Abnormal proliferation, regeneration, degeneration, and atrophy may be  
CC treated by administering the products of the present invention. The  
CC products can also be used in screening assays.  
CC Sequence 1654 BP: 340 A; 551 C; 469 G; 294 T

```

Query Match      1.4%      Score 16; DB 36; Length 1654;
Best Local Similarity 100.0%; Pred. No. 1.86e+01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1433 attcccagagagccca 1448
      |||||
Cp 165 ATTCCCAAGAGGCCA 150

```

```

REUSING 33
ID T7731 standard; DNA; 1809 BP.
AC T7731:
DT 16-SEP-1997 (first entry)
DE DNA encoding 6-hydroxynicotinic acid mono-oxygenase.
KW enzyme, 6-hydroxynicotinic mono-oxygenase, 6-HMO;
P1 16-SEP-1997 (first entry)
P2 16-SEP-1997 (first entry)
OS Pseudomonas fluorescens strain TMS
PS Key Location/Qualifiers
   cda 406..1563
      /tag= 'a'
      /note= '6-hydroxynicotinic acid mono-oxygenase'
   mat_peptide 406..1560
      /tag= 'b'
      /note= 'claim 2'
J09121864.A.
PN 16-SEP-1997
PP 06-NOV-1995; 287441.
PP 06-NOV-1995; JP-287441.
PA (HAEN ) COSMO OIL CO LTD.
PA (USCN ) COSMO SOCO KENTUSHO KK.
PA (USCN ) COSMO SOCO KENTUSHO KK.
P1 16-SEP-1997 (first entry)
P2 16-SEP-1997 (first entry)
P-PDB: W10018.
DR DNA encoding 6-hydroxynicotinic acid mono-oxygenase - useful for

```

```

Matches 16: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 940 attatgcatcaact 955
| | | | | | | | | |
Ov 343 ATCTATGCCATCAACT 358

```

RESULT	35	ID	N90366 standard; cDNA; 2499 BP.
		AC	N90366:1993 (revised)
		CT	13-OCT-1989
		DT	13-OCT-1989 (first entry)
		EW	Synthetic Human 5-lipoxygenase gene.
		KW	Human: 5-lipoxygenase gene; biosynthesis of leukotrienes;
		OS	Homo sapiens (human).
		OS	Homo sapiens (human).
		CD	CDS 35..1962
		FT	/tag= a
		FT	misc_feature 1..34
		FT	/note="attached sequence"
		PN	ZP-325773-A.
		PD	28-AUG-1989.
		PF	28-OCC-1987; 097414.
		PA	Genetic Engineering Center, Inc. 114
		PI	(SISB) Japan Tobacco Inc.
		P1	WT: 89-321819/31.
		P-PDB:	P90730.
		DR	Gene cloning from human 5-lipoxygenase polypeptide
		DR	cDNA clones from human 5-lipoxygenase which catalyzes
		PT	biosynthesis of leukotriene epds
		PT	Clem 1; fig 2; lpp: English.
		PT	Clem 1; fig 2; lpp: English.
		PS	Synthetic human 5-lipoxygenase gene (see corresp. P90730). This
		SS	sequence is identical to the sequence in TGA, YMA or TAG
		SC	is added on, at 1900-1903 the sequence is TGA, YMA or TAG
		QC	sequence 2499 BP: 593 A; 715 C; 537 T; 2 Others:

```

Query Match      1.4%      Score 16; DB 1: Length 7499;
Best Local Similarity 100.0%;
Pred. No. 1.86e+01;
Matches 16; Conservative 0; Indels 0; Gaps 0;
Mismatch 0;
Db 1597 caagctctagcgctc 1612
      |||||

```

QY	995 CAGCTTCAGGCTTC 1011
RESULT	36
Q05178	standard; DNA: 2500 bp.
Q05784	AC
04-JAN-1991	(first entry)
DT	sequence encoding human 5-lipoxygenase.
DE	sequence encoding human 5-lipoxygenase.
KW	E.coli expression vector; leukotriene A4 lipoxins; ds:

	OS	Homologous	Key	Location/Qualifiers
FR	PR	CD	cds	30-12056
FR	PR	CD	/cysg = a	
FR	PR	EP	EP-384750-A.	
FR	PD	PD	29-AUG-1990.	
FR	PP	PP	22-FEB-1990:	301598.
FR	PR	PR	22-FEB-1989:	JP-040206.
FR	PR	(MIS)	(MIS) JAPAN TOBACCO INC.	
FR	PR	WPI	WPI 90-26371/73.	
FR	PR	P-PDS	P-PDS: R05552.	
FR	PR	Prod.	Prod. of human 5-glyoxynase - in large amts. by transformation of E.coli cells, with vector contg. DNA encoding enzyme and culturing.	
FR	PR	Fig	Fig 1; 16pp.	English.
CC	CC	Sequence	Sequence may be used to transform an E.coli expression system for large scale production of human 5-glyoxynase, useful in formation of lektorins and lipoxins.	
CC	CC	Seq	Sequence	2500 BP: 593 A: 706 C: 674 G: 527 T:

```

Query Match      1.44:      Score 16; DB 1; Length 2500;
Best Local Similarity 100.0%;      Pred. No. 1.96e+01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1597 caagctctcaggette 1612
      |||||
Q7 996 cbaagctctcaggettc 1011

```

	RESULT	37	N70805 standard; DNA: 2574 bp.
CC	NC	CC	OJ06951 (first entry)
CC	NC	CC	OJ06951 (last entry)
DE	DE	DE	Sequence of the BglII/XbaI DNA fragment of genomic clone 108-1 of
DE	DE	DE	Eimeria tenella encoding the T4 antigen.
NM	NM	NM	Poultry vaccine: Eimeria tenella vaccine; Eimeria necatrix vaccine;
NM	NM	NM	Coccidiosis; sp.
PF	PF	PF	Eimeria tenella.
PF	PF	PF	Key
PF	PF	PF	Location/Qualifiers
PF	PF	PF	exon
PF	PF	PF	825..912
PF	PF	PF	/tag= a
PF	PF	PF	913..1013
PF	PF	PF	/note= intron 3
PF	PF	PF	/note= intron A'
PF	PF	PF	1014..1189
PF	PF	PF	/tag= c
PF	PF	PF	1190..1303
PF	PF	PF	/note= intron B'
PF	PF	PF	/note= intron B''
PF	PF	PF	1304..1665
PF	PF	PF	exon



FT intron /tag= g  
 FT 1665..1789 /tag= f  
 FT /note= intron C  
 FT 1790..1923 /tag= g  
 PD 04-JUN-1987  
 PP 01-DEC-1986: 065869  
 PR 03-DEC-1985: US-805301.  
 PR 06-DEC-1985: US-805824.  
 PR 11-DEC-1985: US-805824.  
 PR 11-DEC-1985: US-807497.  
 PA (SOLV) SOLWAY & CIE.  
 PI Newman KJ, Gore TC, Tedesco JL, Peterson GR, Brothers VM,  
 PI Kuhn I, McCann MT, Files JG, Sias SR.  
 PI 81-87-199027/23.  
 DR P-PSDB: P70487.  
 DR New polypeptide for protecting chickens against coccidiosis - are  
 FT from recombinant DNA procedures involving total genomic DNA  
 PS Claim 1: Fig 5: 255pp: English.  
 CC Elmeria tenella polypeptides are used to immunise chickens against  
 CC infection against Elmeria strains, so that coccidiosis is prevented.  
 CC They are used in vaccines, e.g. in a dose to provide over 20  
 CC microgram/Ag. The monoclonal antibody or its anti-idiotypic antibody  
 CC may also be administered to confer protection.  
 CC Sequence 2574 BP: 640 A; 594 C; 627 G; 703 T;  
 SQ Query Match 1.48; Score 16; DB 2; Length 2574;  
 Best Local Similarity 100.0%; Pred. No. 1.86e+01; Caps 0;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
 Db 222 ggagattattctggtg 237  
 QY 159 GGGGATATTCGGTG 174  
 RESULT 38  
 IT N70804 standard: DNA: 2575 BP.  
 AC N70804:  
 DT 02-APR-1991 (first entry)  
 DE Sequence of the BglII/EcoRI DNA fragment of genomic clone 108-1 of  
 DE Elmeria tenella encoding the TMA antigen.  
 KW Poultry vaccine; Elmeria tenella vaccine; Elmeria necatrix vaccine;  
 KW as.  
 OS Elmeria tenella.  
 FT key Location/Qualifiers  
 FT 1..914  
 FT /number= 1  
 FT /note= "827..829 = translation initiation codon"  
 FT intron /tag= b  
 FT 915..1015 /tag= c  
 FT 1016..1191 /tag= c  
 FT /number= 2  
 FT intron /tag= e  
 FT 1192..1305 /tag= f  
 FT 1306..1667 /tag= g  
 FT intron /tag= h  
 FT 1668..1789 /tag= i  
 FT /note= "intron C"  
 FT exon /tag= j  
 FT 1790..2424 /tag= k  
 FT /tag= g  
 FT /note= "1925..1927 = termination codon"  
 FT misc\_difference 51  
 FT /tag= h  
 FT /note= "Represents the number 7"  
 FT misc\_difference 57  
 FT /tag= i  
 FT /note= "Represents the number 8"  
 FT misc\_difference 110  
 FT /tag= j  
 FT /note= "Represents the number 5"  
 FT misc\_difference 162  
 FT /tag= k  
 FT /note= "Represents the letter O"  
 FT misc\_difference 177  
 FT /tag= l  
 FT /note= "Printed as Y in the specification"  
 FT misc\_difference 286  
 FT /tag= m  
 FT /note= "Represents the letter O"  
 FT misc\_difference 300  
 FT /tag= n  
 FT /note= "Represents the number 3"

FT intron /tag= b  
 FT 913..1013 /tag= c  
 FT 1014..1189 /tag= c  
 FT intron /tag= d  
 FT 1190..1303 /tag= e  
 FT 1304..1667 /tag= e  
 FT intron /tag= f  
 FT 1668..1789 /tag= g  
 FT 1790..1923 /tag= g  
 PD 04-JUN-1987  
 PP 01-DEC-1986: 065869  
 PR 03-DEC-1985: US-805301.  
 PR 06-DEC-1985: US-805824.  
 PR 11-DEC-1985: US-805824.  
 PR 11-DEC-1985: US-807497.  
 PA (SOLV) SOLWAY & CIE.  
 PI Newman KJ, Gore TC, Tedesco JL, Peterson GR, Brothers VM,  
 PI Kuhn I, McCann MT, Files JG, Sias SR.  
 PI 81-87-199027/23.  
 DR P-PSDB: P70487.  
 DR New polypeptide for protecting chickens against coccidiosis - are  
 FT from recombinant DNA procedures involving total genomic DNA  
 PS Claim 1: Fig 5: 255pp: English.  
 CC Elmeria tenella polypeptides are used to immunise chickens against  
 CC infection against Elmeria strains, so that coccidiosis is prevented.  
 CC They are used in vaccines, e.g. in a dose to provide over 20  
 CC microgram/Ag. The monoclonal antibody or its anti-idiotypic antibody  
 CC may also be administered to confer protection.  
 CC Sequence 2575 BP: 640 A; 601 C; 624 G; 706 T;  
 SQ Query Match 1.48; Score 16; DB 2; Length 2575;  
 Best Local Similarity 100.0%; Pred. No. 1.86e+01; Caps 0;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
 Db 222 ggagattattctggtg 237  
 QY 159 GGGGATATTCGGTG 174  
 RESULT 39

ID Q37481 standard: DNA: 2577 BP.  
 AC Q37481:  
 DT 11-JUN-1993 (first entry)  
 DE Sequence of the BglII/EcoRI DNA fragment of genomic clone 108-1 of  
 DE Elmeria tenella encoding the TMA antigen.  
 KW Poultry vaccine; Elmeria tenella vaccine; Elmeria necatrix vaccine;  
 KW as.  
 OS Elmeria tenella.  
 FT key Location/Qualifiers  
 FT 1..914  
 FT /number= 1  
 FT /note= "827..829 = translation initiation codon"  
 FT intron /tag= b  
 FT 915..1015 /tag= c  
 FT 1016..1191 /tag= c  
 FT /number= 2  
 FT intron /tag= e  
 FT 1192..1305 /tag= f  
 FT 1306..1667 /tag= g  
 FT intron /tag= h  
 FT 1668..1789 /tag= i  
 FT /note= "intron C"  
 FT exon /tag= j  
 FT 1790..2424 /tag= k  
 FT /tag= g  
 FT /note= "1925..1927 = termination codon"  
 FT misc\_difference 51  
 FT /tag= h  
 FT /note= "Represents the number 7"  
 FT misc\_difference 57  
 FT /tag= i  
 FT /note= "Represents the number 8"  
 FT misc\_difference 110  
 FT /tag= j  
 FT /note= "Represents the number 5"  
 FT misc\_difference 162  
 FT /tag= k  
 FT /note= "Represents the letter O"  
 FT misc\_difference 177  
 FT /tag= l  
 FT /note= "Printed as Y in the specification"  
 FT misc\_difference 286  
 FT /tag= m  
 FT /note= "Represents the letter O"  
 FT misc\_difference 300  
 FT /tag= n  
 FT /note= "Represents the number 3"

FT misc\_difference 427  
 FT /tag= o  
 FT /note= "Printed as Y in the specification"  
 FT misc\_difference 451  
 FT /tag= p  
 FT /note= "Printed as M in the specification"  
 FT misc\_difference 514  
 FT /tag= q  
 FT /note= "Printed as M in the specification"  
 FT misc\_difference 615  
 FT /tag= r  
 FT /note= "Represents the number 7"  
 FT misc\_difference 673  
 FT /tag= s  
 FT /note= "Represents the number 7"  
 FT misc\_difference 178  
 FT /tag= t  
 FT /note= "Represents the number 7"  
 PD 04-JUN-1987  
 PP 01-DEC-1986: 065869  
 PR 03-DEC-1985: US-805301.  
 PR 06-DEC-1985: US-805824.  
 PR 11-DEC-1985: US-805824.  
 PR 11-DEC-1985: US-807497.  
 PA (SOLV) SOLWAY & CIE.  
 PI Newman KJ, Gore TC, Tedesco JL, Peterson GR, Brothers VM,  
 PI Kuhn I, McCann MT, Files JG, Sias SR.  
 PI 81-87-199027/23.  
 DR P-PSDB: P70487.  
 DR Nucleic acid encoding Elmeria tenella proteins - capable of  
 FT inducing immune response in chickens, useful in vaccine  
 PS Claim 1: Fig 5: 255pp: English. DNA isolated from Elmeria  
 CC apiculatus oocysts. The library was screened with oligonucleotide  
 CC probes based on peptide fragments of the purified 17,000 Dalton  
 CC component of the TMA antigen. Clone 108-1 hybridised strongly to  
 CC one probe and weakly to the others. This positive clone was  
 CC subcloned and sequenced. The sequence of the clone was  
 CC compared with the sequence of the TMA antigen. It encoded a 21 amino acid  
 CC signal peptide which was removed from the mature, purified protein.  
 CC The DNA also contained 3 introns. From the predicted amino acid  
 CC sequence of the mature protein, it seems that the two  
 CC peptide of the TMA antigen (17 kD and 10 kD) are  
 CC contiguous nucleotide sequence, and at least one proteolytic step  
 CC occurs to generate the 8 kD peptide by removal of a tripeptide.  
 CC The sequence as printed in the specification includes several  
 CC non-coded amino acids (underlined) which have been replaced by N  
 CC for inclusion in the Genbank record.  
 SQ Sequence 2577 BP: 640 A; 594 C; 627 G; 703 T;











```

CC CC gene, somatic mutations of which in the region spanning nucleotides
CC 3805-4889 of the NF1 CDM, in human tumours, indicates defective
CC ras regulation. Therefore a tumor found to contain a somatic
CC mutation in the NF1 gene can be treated using ras activity as the
CC basis for treatment. A tumor containing a somatic mutation
CC other courses of treatment. A tumor containing a somatic mutation
CC in the NF1 gene can be treated by inactivating ras p21, also as GAP
CC p120 is present, but apparently latent, GAP p120 activation would
CC counteract the loss of HNF or hnf GAP related domain activity, and also
CC counteract the loss of hnf or hnf GAP related domain activity.
CC SQ Sequence 10706 BP; 3993 A; 2454 C; 2410 G; 2849 T;
CC
CC Query Match 1.48; Score 16; DP 29; Length 10706;
CC Best Local Similarity 100.0%; Prod No 1866+0; Indels 0; Gaps 0;
CC Matches 16; Conservative 0; Mismatches 0;
CC
Dd 801 ttgatgttcattgatat 816
Cc 652 tttatgttcatgatat 667
CC
RESID 50
CC TT2460 standard; DNA: 2954 BP.
AC AC 121460;
AD DT 21-APR-1996 (first entry)
AE Ovine adenovirus genome.
AF E; promoter; packaging signal; RNA processing signal; ELA/B;
AK E; promoter; packaging signal; RNA processing signal;
AL plasmid pOA100; cloning; vector; antigen; recombinant vaccine;
AW rotavirus; Trichostonylus colubriformis; Taenia ovis;
AX gene transfer; cyclic; RNA; ribosome; transgenic animal;
AY gene transfer; cyclic; RNA; ribosome; transgenic animal;
AZ Ovine adenovirus strain OA2187.
BA K Location/Qualifiers
BB repeat_region ..46
BB ..46
BB /rpt_type= INVERTED
BB /note= "Inverted terminal repeat (claim 7)."
BT misc_feature 21139..22130
BT ..46
BT /note= "Non-essential region suitable for cloning."
BU misc_feature 26645..26646
BU ..c
BU /tag= c
BV misc_feature 26645..26646
BV ..c
BW misc_feature 26645..26646
BW ..c
BX misc_feature 26645..26646
BX ..c
BY misc_feature 26645..26646
BY ..c
BZ misc_feature 26645..26646
BZ ..c
CA MO9603508.A1.
CB NC_007101.1.
CD 26-JUN-1995; ADU0453
CE 26-JUL-1995; ADU07101.
CF PR

```

Tue Nov 17 08:55:26 1998

US-08-887-977-9. 109

PI	Calbretta B, Szowski T;
D	WFI: 98-27988(1);
PT	Anticancer composition comprising two anti-sense oligo:nucleotides(s)
PT	- targeting cytoplastic and nuclear oncogene(s)
PT	Claim 1; Column 55-59; 9pp; English.
CC	The present invention describes a composition which comprises two
CC	antisense oligonucleotides. The first oligonucleotide is specific for a
CC	cytoplasmic oncogene or proto-oncogene selected from ras, raf, MyB-1,
CC	c-myc, c-fos, c-jun, c-abl, bcr-abl, v-src, v-lsrc, v-hsrc, v-msrc,
CC	cyes'. The second oligonucleotide is specific for a nuclear oncogene or
CC	proto-oncogene selected from myc, jun, c-ras, c-fos, c-myb, c-mysb,
CC	c-rel, c-zav, c-k1, c-spl, cyclin D1, PML/RAR alpha , AML1/APMG,
CC	TALP1 and ABL-2/APMG. The composition is used for treating cancer.
CC	The composition can be administered intravenously, intraperitoneally
CC	or orally. The composition has been shown to synergistically
CC	enhanced ability to inhibit growth of cancer cells.
SQ	Sequence 35100 BP: 847A A: 8597 C: 9682 G: 8347 T:
	Query Match 1.48; Score 15; DB 40; Length 35100;
	Best Local Similarity 100.0%; Prd No. 1.86e+01;
	Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	26273 catcacacacagaga 26289
CP	488 CACACACAGGACGA 473

[illegible][illegible]

Page 72

IS-08-887-977-9. rna

[illegible]







```

CC other 8 creat an Shoi site 5' to the start codon. Primer 2 is
CC complementary to an internal part of the coding region, contg. a
CC Balli site which is present in the gene itself. These 2 primers
CC permit the amplification of a 335 base fragment.
SQ Sequence 22 BP; 6 A.; 5 C.; 0 G.; 0 T.; 0 Mismatches 0; Indels 0; Gaps 0;
Query Match 1.3% Score 14; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. NO. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 9 attcaagattctca 22
|||||
Cc 969 ctttcacattctca 956

RESULT 55
ID LQ86004 standard; DNA; 24 BP.
DB 8-NOV-1995 (first entry)
DE Brevibacterium flavum plasmid pOOP 1 PCR primer 2.
KW Brevibacterium flavum; plasmid pOOP 1; coryneform bacteria;
OS Streptococcaceae; recombinant production; PCR primer 2; as.
PW J07033383-A.
PR 10-FEB-1995. 159857
PF 28-JUN-1993. 375957
PI 28-JUN-1993. 375957
PA (MTP.) NITSUBISHI PETROCHEMICAL CO LTD.
DE W21: 95-117862/16.
PT Plasmid vector for transformation of coryneform bacteria - used
PT for synthesis of useful substances
PE Synthesized by P.D. 172170
CC Q86013 and Q86004 are a pair of primers for the PCR amplification
CC of Q86002, which encodes R71617 the brevibacterium flavum plasmid
CC pOOP 2 protein product. The DNA was used in the construction of a
CC coryneform bacterium expressing of the E. coli beta galactosidase in
CC coryneform bacteria.
SQ Sequence 24 BP; 5 A.; 9 C.; 7 G.; 0 T.; 0 Mismatches 0; Indels 0; Gaps 0;
Query Match 1.3% Score 14; DB 15; Length 24;
Best Local Similarity 100.0%; Pred. NO. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 5 tccgcgcagcaca 18
|||
Cv 455 TACCACGACGACA 468

RESULT 56
ID O83217 standard; DNA; 24 BP.
DB 8-OCT-1995 (first entry)
DE Brevibacterium flavum plasmid pOOP 1 PCR primer 2.
KW Brevibacterium flavum; plasmid pOOP 1; coryneform bacteria;
OS Streptococcaceae; recombinant production; PCR primer 2; as.
PW J07033383-A.
PR 10-FEB-1995. 159857
PF 28-JUN-1993. 375957
PI 28-JUN-1993. 375957
PA (MTP.) NITSUBISHI PETROCHEMICAL CO LTD.
DE W21: 95-117862/16.
PT Plasmid vector for transformation of coryneform bacteria - used
PT for synthesis of useful substances
PE Synthesized by P.D. 172170
CC Q86013 and Q86004 are a pair of primers for the PCR amplification
CC of Q86002, which encodes R71617 the brevibacterium flavum plasmid
CC pOOP 2 protein product. The DNA was used in the construction of a
CC coryneform bacterium expressing of the E. coli beta galactosidase in
CC coryneform bacteria.
SQ Sequence 24 BP; 5 A.; 9 C.; 7 G.; 0 T.; 0 Mismatches 0; Indels 0; Gaps 0;
Query Match 1.3% Score 14; DB 15; Length 24;
Best Local Similarity 100.0%; Pred. NO. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 5 tccgcgcagcaca 18
|||
Cv 455 TACCACGACGACA 468

```

Tue Nov 17 08:55:26 1998

This primer is used to synthesise the 3' end of the human Col XII gene by PCR amplification. This is used in the construction of recombinant vectors containing collagen genes. A novel method for producing a (pro)collagen polypeptide comprises culturing a host cell, where the host cell contains a transfected plasmid, with a polynucleotide having a first expression vector which encodes a polynucleotide subunit and a second acid sequence which encodes a (pro)collagen subunit and a second expression vector comprising a polynucleotide molecule having a nucleic acid sequence which encodes at least one (pro)collagen polypeptide transmembrane domain. The method provides for the synthesis of correctly folded proteins so that they exhibit the normal triple-helical conformation characteristic of procollagens and collagens. Purification of the collagens is optionally facilitated.

Sequence 38 BP: 12 A C; 11 G; 5 T;

Query Match 1-38; Score 15.38; Length 38;  
Best Local Similarity 100.0%; Read No. 8378-01;  
Matrix 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 cttctataaagcaaa 31  
Cp 190 CTTCTATAAAGCAAA 184  
|||||

RESULT 58

AD 025332 standard: DNA; 40 BP.  
ID 025332 standard: DNA; 40 BP.  
LN 025332 standard: DNA; 40 BP.  
DE 06-DE-1994 (first entry)  
DE Oligonucleotide probe for repeated sequence of D17X1 locus.  
KW Probe; detection; repeat sequence; D17X1 locus; control region;  
KW Subunit sequence; human alpha satellite locus; mitochondria; ss.  
PR 58-594959-A.  
PF 04-MAY-1994.  
PD 04-MAY-1994.  
PI 132661  
PF 28-NOV-1993; 132661  
PI 28-NOV-1993; 132661  
PA (HOFF) JOFFMAN LA ROCHE & CO AG F.  
PI Reynolds RL, Walsh PS;  
PI WPI; 94:145677/18.  
PI oligonucleotide probes for detecting human DNA - having  
PI conserved sequence in the mitochondrial genome or a  
PI conserved sequence in the mitochondrial genome  
PI Claim 6; Page 13; 27pp; English.  
CC This oligonucleotide probe is complementary to a region in the 2,7  
CC probe complementary to a highly repeated sequence or to a sequenced

Page 79

This primer is used to synthesise the 3' end of the human Col. III gene by PCR amplification. This is used in the construction of recombinant vectors containing collagen genes. A novel method for producing a (pro)collagen polypeptide comprises culturing a host cell, where the host cell has been transfected with DNA encoding the first chain of the expression vector coding for procollagen type I alpha 1(I), having nucleic acid sequence which encodes a (pro)collagen subunit and a second expression vector comprising a polynucleotide molecule having a nucleic acid sequence which encodes at least one (pro)collagen post-translational modification site; growing said cells under conditions such as to produce from the cultured cell. The (pro)collagen polypeptide is selected from collagen types IV, V, VI, VII, VIII, IX, X, XI, XII, XIII, XIV, XV, XVI, XVII, XVIII, and XIX. The methods can be used for the production of cDNA libraries or other biological samples. The invention also includes applications. The method provides for the synthesis of correctly folded proteins so that they exhibit the normal triple-helical conformation characteristic of procollagens and collagens. Purification of the collagen is greatly facilitated.

Query Match  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DG 17 ctctaaagaacaaa 31  
CG 198 CTTATAAAGAACAAA 184

RESULT 58  
ID Q62332 standard; DNM; 40 BP.  
LN LN  
DT 06-DEC-1994 (first entry)  
DE Oligonucleotide probe for repeated sequences of DI7k1 locus.  
KW Probe; detection; repeat sequence; DI7k1; locus; control region;  
OS Synthesized sequence; Human Alpha satellite locus; mitochondria; ss.  
PN EP-594959-A.  
PD 04-MAY-1994.  
PT 28-AUG-1993; 133281.  
RF JOURNAL OF MOLECULAR EVOLUTION  
PA ((ROFF)) HOFFMANN LA ROCHE & CO AG F.  
PI Reynolds RL, Walsh PS:  
WPI: 94-145673/18.  
DR Oligonucleotide probes for detecting human DNA - having  
SA conserved sequence in the mitochondrial genome  
PS conserved sequence in the mitochondrial genome  
FI Claim 6; Page 13; 27pp; English.  
CC This oligonucleotide probe is complementary to a region in the 2.7  
CC kilobase pair deletion product described above.  
CC Proben complementary to a highly-repeated sequence or to a conserved

13-NDV-1995 (first entry)  
 DE Drevibacterium flavum M2333 plasmid pCOP1 and pCOP2 PCR primer 2.  
 RW Drevibacterium flavum: M2333; plasmid pCOP1 and pCOP2;  
 KW chloramphenicol resistant; stably maintained; PCR primer 2; se.  
 PD 370338192-A  
 PD 370338192-A  
 PD 10-FEB-1995  
 PP 28-JUN-1993; 150748  
 PR 28-JUN-1993; 150748  
 PR 28-JUN-1993; 150748  
 DR WPI: 95-117861/16  
 DR PIETROCHEMICAL CO LTD.  
 DR Plasmid vector stably maintained in Corynebacterium - useful  
 PT for the expression of genes of interest in coryneform bacteria  
 CC O38215 and O38217 are pair of primers for the PCR amplification  
 CC O38216 and O38217 are pair of primers for the PCR amplification  
 CC of O38215, the Streptibacterium flavum M2333 plasmid pCOP1 and pCOP2,  
 CC which encodes R0143 a chloramphenicol resistance protein. This  
 CC plasmid is stably maintained in the host and can be inherited  
 CC plasmid is stably maintained in the host and can be inherited  
 CC Sequence 9758201/48 5 A; 9 C; 7 G; 3 T;  
 SQ Sequence 24 BP: 138; Score 14; DB 15; Length 24;  
 Query Match: Similarity 100%; Read M; 370702;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 5 tacgcgcgcgcacaa 18  
 Oy 455 TACCGCGCGCGACAA 48

RESULTS 57  
 AC V06323; standard; DNA; 38 BP.  
 AC V06323;  
 DE 06-MAY-1998 (first entry)  
 DE Human Col III gene 3' end synthesizing 3' primer.  
 DE Human Col III gene 3' end synthesizing 3' primer.  
 DE Human Col III gene 3' end synthesizing 3' primer.  
 OS procollagen; PCR primer; se.  
 OS Synthetic.  
 OS Homo sapiens.  
 FN 7097387/1957  
 FN 7097387/1957  
 PP 11-APR-1997; 007300.  
 PP 12-APR-1996; 09-631336.  
 PA (PIET.) ACAD FINLAND.  
 PA PIETROCHEMICAL CO LTD.  
 PA Kivirikki KJ, Pihlajaniemi T.  
 WPI: 97-526201/48  
 DR Recombinant production of (pro)collagen having correct folding -  
 DR using vectors encoding collagen sub-unit and collagen  
 DR using vectors encoding collagen sub-unit and collagen  
 PS Example 11; Pages 67; 90pp; English.

Nov 17 08:55:26 1998 US-08-887-977-9. rmg

CC sequence provides high detection sensitivity and specificity. The  
CC peptide was particularly effective when used in combination with  
CC quantitative the amount of DNA in a sample. See 06352-51.  
SC Sequence 40 BP; 12 A; 7 C; 7 G; 14 T;  
Query Match 1.3%; Score 14; DB 11; Length 40;  
Best Local Similarity 100.0%; Pred. No. 3,476+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 12 KGAAGAACTACTTT 25  
QY 941 TGAAGAATCACTTT 954  
RESULT 59  
ID T07129 standard; cDNA; 48 BP.  
DT 701729  
DT 26-JUN-1996 (first entry)  
KW Linguar antimicrobial peptide; epithelium; LAP; bovine; beta-defensin;  
KW antimicrobial activity; Gram-positive bacteria; Gram-negative bacteria;  
KW fungal pathogen; mammal; microbial infection; immunodeficiency; AIDS;  
KW Cystic fibrosis; gum disease; burn; pneumonia; human; probe; ss.  
PR W59532287-A1.  
PR 30-MAY-1995.  
PR 24-NOV-1995. 006761s.  
PR 24-NOV-1995. 006761s.  
PR (MAG.) MAGALIN PHARM INC.  
PI Schönwetter BS, Zaslloff MA;  
DR WPI: 86-020587/02.  
CC Treating cutaneous fungal infections with the epithelium,  
CC e.g. in gum disease, cystic fibrosis, burns etc.  
PR Diagnostics; Page 5; 37pp; English.  
CC 48-mer probe. LAP is a member of the beta-defensin group of the defensin  
CC family of peptides. LAP has broad spectrum antimicrobial activity  
CC against Gram-positive and Gram-negative bacteria, and fungal pathogens.  
CC It is present at high levels in mammalian epithelia, with high expression  
CC levels induced by various agents. The LAP gene can be used as a method of  
CC cDNA encoding endogenous up-regulators of LAP. In this method, the levels of  
CC cells are cultured in the presence of a test substance. The epithelial  
CC up-regulator. LAP is used to treat microbial infections of the  
CC epithelium (or of those that extend to deeper tissues) e.g. in  
CC immunodeficient states (AIDS), cystic fibrosis, gum disease, wounds,  
CC burns and pneumonia. 8 A; 15 C; 11 G; 14 T.



Query Match 1.38; Score 15; DB 18; Length 48;  
Best Local Similarity 100.00; Pred. No. 8.37e+01; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 20 cggcctcagcag 34  
Cp 595 TGGGCTCCAGCAG 581

RESULT 60  
ID G62540 standard; DNA: 60 BP.  
AC G62540.  
DT 06-DEC-1994 (first entry)  
DE Oligonucleotide probe for repeated sequence of D17S1 locus.  
KW probe detection; repeat sequence; D17S1 locus; control region;  
KW Synthetic sequence; human alpha satellite locus; mitochondria; ss.  
OS Synthetic.  
PN EP-594959-A.  
PD 04-MAY-1994. 113961.  
PR 28-AUG-1992. DS-938084.  
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.  
PI Reynolds RL, Walsh PS;  
DR WPI: 94-15577/18.  
PT Oligonucleotide probes for detecting human DNA - having  
PT conserved sequence in the mitochondrial genome  
PS Claim 6; Page 17; 27pp; English.  
CC This oligonucleotide probe is complementary to a region in the 2.7  
CC kilobase region of the mitochondrial genome. The use of  
CC probes complementary to highly repeated sequence, or highly  
CC conserved sequence provides high detection sensitivity and specificity. The  
CC probe may be optionally labelled with biotin and used in methods to  
CC quantitate the amount of DNA in a sample. See 062532-51.  
SQ Sequence 60 BP; 20 A; 10 C; 11 G; 19 T;

Query Match 1.38; Score 14; DB 11; Length 60;  
Best Local Similarity 100.00; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 12 tcgaactactct 25  
Cp 941 TCAGAACTACTCTT 954

RESULT 61  
ID T28088 standard; DNA: 72 BP.  
AC T28088.  
DT 11-DEC-1996 (first entry)  
DE Probe OGC3 isolated from fibroblasts.

Polymerase chain reaction; PCR: primer: amplify; human; fibroblast; AIDS;  
enhanced differential display; EDD; mRNA preparation; senescent cell;  
KW enhanced differential display; EDD; mRNA preparation; senescent cell;  
KW non-senescent cell; aged; related lipofuscin accumulation; gene expression;  
KW donor tissue; senescent melanocyte; melanin; hypopigmentation; ss.  
OS Synthetic.  
PN W09613610-A2.  
PR 24-AUG-1995; U11230.  
PA (GERO-) GERON CORP.  
PI Feng W, Hirsch KS, Lindeken MK, Villeponteau B;  
DR WPI: 95-251464/25.  
PT Identifying, isolating and regulating senescence-related genes -  
PT useful to ameliorate problems associated with accumulation of  
PT senescent cells, e.g. age-related lipofuscin accumulation in the  
PT donor tissue.  
PS Claim 8; Page 36; 135pp; English.  
CC T28076-T28113, and T28131-T28173 represent novel senescent-related gene  
CC sequences isolated from fibroblasts using the method of the invention.  
CC in the method of the invention, mRNA was isolated from a senescent cell,  
CC such as those shown in T28044-T28075 in separate reaction mixtures. The  
CC amplified sequences are then separated by size or charge, and the  
CC products are analyzed to identify a gene from young quiescent cells and  
CC cDNA library that is present at a different level from senescent  
CC cells. The method can be used to identify genes that are up-regulated  
CC in conjunction with an enhanced differential display (EDD) method (an mRNA  
CC preparation method) on the fibroblasts. The method can be used for the  
CC rapid and efficient identification and isolation of senescence-related  
CC genes and non-senescent cells. It can also be used to destroy cells expressing  
CC senescence specific (or related) gene products, and to screen for  
CC compounds capable of altering gene expression in senescent cells. The  
CC method can also be used to ameliorate problems associated with the  
CC accumulation of senescent cells in the treatment of AIDS. Also, the  
CC method can be used to distinguish young cells from senescent cells in  
CC donor tissue, which is useful in removing senescent melanocytes  
CC overexpressing melanin which cause hypopigmentation, or liver spots.  
SQ Sequence 72 BP; 34 A; 13 G; 14 C; 15 T;

Query Match 1.38; Score 14; DB 24; Length 72;  
Best Local Similarity 100.00; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 35 agactctcaag 48  
Cp 707 AGAATCTCAAGG 720

Query Match 1.38; Score 14; DB 22; Length 78;  
Best Local Similarity 100.00; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 65 tctatataagcaga 78  
Cp 157 TTATATAAGCAGA 184

RESULT 64  
ID G62541 standard; DNA: 80 BP.  
AC G62541.  
DT 06-DEC-1994 (first entry)  
DE Oligonucleotide probe for repeated sequence of D17S1 locus.  
KW probe detection; repeat sequence; D17S1 locus; control region;  
KW Synthetic sequence; human alpha satellite locus; mitochondria; ss.  
OS Synthetic.

Gene signature, messenger RNA, mRNA, relative abundance, frequency;  
human; cloning; mapping; non-biased library; diagnosis; detection;  
KW cell typing; abnormal cell function; ss.  
PN W0951472-A1.  
PR 11-NOV-1994; J01916.  
PA (MATSU) MATSUOKA K.  
PI (KUMU) KUMU K.  
DR WPI: 95-206931/27.  
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT cells.  
PS Claim 1; Page 1745; 2245pp; Japanese.  
CC A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
CC derived from a human cDNA library, and which is used for the diagnosis of  
CC human genomic DNA cDNA or mRNA is claimed. The GS (Gene Signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(rI) as the sole primer. Since the 3'-  
CC end of the cDNA is complementary to the 5'-end of the mRNA, each  
CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined by hybridizing the library with a specific mRNA (or its  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognizing different cell types.  
SQ Sequence 78 BP; 26 A; 10 C; 14 G; 26 T;

Query Match 1.38; Score 14; DB 22; Length 78;  
Best Local Similarity 100.00; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ds 65 tctatataagcaga 78  
Cp 157 TTATATAAGCAGA 184

RESULT 64  
ID G62541 standard; DNA: 80 BP.  
AC G62541.  
DT 06-DEC-1994 (first entry)  
DE Oligonucleotide probe for repeated sequence of D17S1 locus.  
KW probe detection; repeat sequence; D17S1 locus; control region;  
KW Synthetic sequence; human alpha satellite locus; mitochondria; ss.  
OS Synthetic.



PN EP-594955-A.  
 PD 04-MAY-1994. 113261.  
 PR 28-NOV-1993; US-938084.  
 PA (HOFF) ROFFMAN LA ROCHE & CO AG F.  
 PI Reynolds RL, Walsh PS;  
 DR 97: 94-15673/18. Probes for detecting human DNA - having  
 PT conserved sequence in the mitochondrial genome  
 CC complementary to a human alpha satellite locus or a  
 PS Claim 6; Page 17; 27pp; English.  
 CC This oligonucleotide probe is complementary to a region in the 2.7  
 CC kilobase D17S1 locus, a highly repetitive sequence. The use of  
 CC probes complementary to a highly repeated sequence or to a conserved  
 CC sequence provides high detection sensitivity and specificity. The  
 CC probe may be optionally labelled with biotin and used in methods to  
 CC quantitate the amount of DNA in a sample. See 06522-51.  
 CC Sequence 80 BP; 24 A; 13 G; 16 G; 25 T;  
 CC Sequence 80 BP; 24 A; 13 G; 16 G; 25 T;  
 Query Match 1.38; Score 14; DB 11; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 tcgaactactctt 25  
 QY 941 TCAGAGACTACTCTT 954

RESULT 65  
 ID 724568 standard; cDNA to mRNA; 80 BP.  
 AC 054688; 113261.  
 DE Human gene signature HUMG06618.  
 KW Gene signature: messenger RNA; mRNA; relative abundance; frequency;  
 KW human; cloning; mapping; non-biased library; diagnosis; detection;  
 OS Homo sapiens; abnormal cell function; ss.  
 PN M09514772-A1.  
 PD 01-JUN-1995.  
 PR 11-NOV-1994; J01916504.  
 PA (MATSU) MATSUBARA K.  
 PI (OKUBO) OKUBO K.  
 PI Matsubara K, Okubo K;  
 PR 74: 12-706931/27. Probes for detecting human DNA - having  
 PT conserved sequence in the mitochondrial genome  
 CC complementary to a human alpha satellite locus or a  
 PS Claim 6; Page 17; 27pp; English.  
 CC This oligonucleotide probe is complementary to a region in the 2.7  
 CC kilobase D17S1 locus, a highly repetitive sequence. The use of  
 CC probes complementary to a highly repeated sequence or to a conserved  
 CC sequence provides high detection sensitivity and specificity. The  
 CC probe may be optionally labelled with biotin and used in methods to  
 CC quantitate the amount of DNA in a sample. See 06522-51.  
 CC Sequence 80 BP; 24 A; 13 G; 16 G; 25 T;  
 CC Sequence 80 BP; 24 A; 13 G; 16 G; 25 T;  
 Query Match 1.38; Score 14; DB 11; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 tcgaactactctt 25  
 QY 941 TCAGAGACTACTCTT 954

CC double-stranded DNA) which comprises one of the 7837 'GS' sequences  
 CC given in T19001-726837 and which is able to hybridize to part of  
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
 CC is a double-stranded DNA sequence which is able to hybridize to part of  
 CC from various human tissues; synthesis of cDNA was initiated from the  
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence is unique to a particular mRNA species, almost  
 CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library  
 CC contains a different mRNA in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC determined (esp. using primers and probes derived from the GS  
 CC sequence) as a means of diagnosing abnormal cell function or for  
 CC recombinant different cell types.  
 CC Sequence 80 BP; 33 A; 9 C; 9 G; 29 T;  
 Query Match 1.38; Score 15; DB 21; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-01;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 23 tgaacaaataatca 36  
 Cp 676 TGAACAAATAATCA 662

RESULT 66  
 ID 054688 standard; cDNA; 95 BP.  
 AC 054688; 113261.  
 DE Another specific cDNA clone ant52.  
 KW Transgenic plants; male sterility; pollen; sterile;  
 KW CytA; toxin; Nicotiana tabacum; Diptheria; Erwinia chrysanthemi;  
 KW phase Mu; Pseudomonas syringae; Bacillus thuringiensis; anther;  
 KW breeding; ss.  
 OS Nicotiana tabacum  
 PN M09514772-A1.  
 PD 01-JUN-1995.  
 PR 11-NOV-1994; J01916504.  
 PA (MATSU) MATSUBARA K.  
 PI (OKUBO) OKUBO K.  
 PI Matsubara K, Okubo K;  
 PR 74: 12-706931/27. Probes for detecting human DNA - having  
 PT conserved sequence in the mitochondrial genome  
 CC complementary to a human alpha satellite locus or a  
 PS Claim 6; Page 17; 27pp; English.  
 CC This oligonucleotide probe is complementary to a region in the 2.7  
 CC kilobase D17S1 locus, a highly repetitive sequence. The use of  
 CC probes complementary to a highly repeated sequence or to a conserved  
 CC sequence provides high detection sensitivity and specificity. The  
 CC probe may be optionally labelled with biotin and used in methods to  
 CC quantitate the amount of DNA in a sample. See 06522-51.  
 CC Sequence 80 BP; 33 A; 9 C; 9 G; 29 T;  
 Query Match 1.38; Score 15; DB 21; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-01;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 23 tgaacaaataatca 36  
 Cp 676 TGAACAAATAATCA 662

CC Another specific cDNA or genomic sequences can be used to identify  
 CC and isolate another specific promoters. The anther specific promoter  
 CC can then be cloned into a recombinant construct and used to express  
 CC heterologous genes. Preferred heterologous genes include Diptheria  
 CC toxin, Nicotiana tabacum, Diptheria, Erwinia chrysanthemi,  
 CC chrysanthem; T-urif3 from cma-T maize mitochondrial genomes; the  
 CC Gln recombinase gene from phase Mu; the indole acetic acid-lysin  
 CC synthetase gene from Pseudomonas syringae and the CytA toxin gene  
 CC from Bacillus thuringiensis. In all of these genes, the 3'-end of  
 CC the mRNA is unique to a particular mRNA species, almost all the  
 CC 3'-oriented cDNAs hybridize with specific mRNAs. Each library  
 CC contains a different mRNA in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC determined (esp. using primers and probes derived from the GS  
 CC sequence) as a means of diagnosing abnormal cell function or for  
 CC recombinant different cell types.  
 CC Sequence 110 BP; 33 A; 22 C; 12 G; 37 T;  
 Query Match 1.38; Score 14; DB 20; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 21 actctaaagcca 34  
 QY 709 ATTCTAAAGCCA 722

RESULT 67  
 ID 054688 standard; DNA; 100 BP.  
 AC 054688; 113261.  
 DE Human gene signature HUMG06618.  
 KW Gene signature: messenger RNA; mRNA; relative abundance; frequency;  
 KW human; cloning; mapping; non-biased library; diagnosis; detection;  
 OS Homo sapiens; abnormal cell function; ss.  
 PN M09514772-A1.  
 PD 01-JUN-1995.  
 PR 11-NOV-1994; J01916504.  
 PA (MATSU) MATSUBARA K.  
 PI (OKUBO) OKUBO K.  
 PI Matsubara K, Okubo K;  
 PR 74: 12-706931/27. Probes for detecting human DNA - having  
 PT conserved sequence in the mitochondrial genome  
 CC complementary to a human alpha satellite locus or a  
 PS Claim 6; Page 17; 27pp; English.  
 CC This oligonucleotide probe is complementary to a region in the 2.7  
 CC kilobase D17S1 locus, a highly repetitive sequence. The use of  
 CC probes complementary to a highly repeated sequence or to a conserved  
 CC sequence provides high detection sensitivity and specificity. The  
 CC probe may be optionally labelled with biotin and used in methods to  
 CC quantitate the amount of DNA in a sample. See 06522-51.  
 CC Sequence 110 BP; 33 A; 22 C; 12 G; 37 T;  
 Query Match 1.38; Score 14; DB 20; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 21 actctaaagcca 34  
 QY 709 ATTCTAAAGCCA 722

CC Another specific cDNA or genomic sequences can be used to identify  
 CC and isolate another specific promoters. The anther specific promoter  
 CC can then be cloned into a recombinant construct and used to express  
 CC heterologous genes. Preferred heterologous genes include Diptheria  
 CC toxin, Nicotiana tabacum, Diptheria, Erwinia chrysanthemi,  
 CC chrysanthem; T-urif3 from cma-T maize mitochondrial genomes; the  
 CC Gln recombinase gene from phase Mu; the indole acetic acid-lysin  
 CC synthetase gene from Pseudomonas syringae and the CytA toxin gene  
 CC from Bacillus thuringiensis. In all of these genes, the 3'-end of  
 CC the mRNA is unique to a particular mRNA species, almost all the  
 CC 3'-oriented cDNAs hybridize with specific mRNAs. Each library  
 CC contains a different mRNA in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC determined (esp. using primers and probes derived from the GS  
 CC sequence) as a means of diagnosing abnormal cell function or for  
 CC recombinant different cell types.  
 CC Sequence 110 BP; 33 A; 22 C; 12 G; 37 T;  
 Query Match 1.38; Score 14; DB 20; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 tcgaactactctt 25  
 QY 941 TCAGAGACTACTCTT 954

RESULT 68  
 ID 722661 standard; cDNA to mRNA; 110 BP.  
 AC 054688; 113261.  
 DE Human gene signature HUMG06618.  
 KW Gene signature: messenger RNA; mRNA; relative abundance; frequency;  
 KW human; cloning; mapping; non-biased library; diagnosis; detection;  
 OS Homo sapiens; abnormal cell function; ss.  
 PN M09514772-A1.  
 PD 01-JUN-1995.  
 PR 11-NOV-1994; J01916504.  
 PA (MATSU) MATSUBARA K.  
 PI (OKUBO) OKUBO K.  
 PI Matsubara K, Okubo K;  
 PR 74: 12-706931/27. Probes for detecting human DNA - having  
 PT conserved sequence in the mitochondrial genome  
 CC complementary to a human alpha satellite locus or a  
 PS Claim 6; Page 17; 27pp; English.  
 CC This oligonucleotide probe is complementary to a region in the 2.7  
 CC kilobase D17S1 locus, a highly repetitive sequence. The use of  
 CC probes complementary to a highly repeated sequence or to a conserved  
 CC sequence provides high detection sensitivity and specificity. The  
 CC probe may be optionally labelled with biotin and used in methods to  
 CC quantitate the amount of DNA in a sample. See 06522-51.  
 CC Sequence 110 BP; 33 A; 22 C; 12 G; 37 T;  
 Query Match 1.38; Score 14; DB 20; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 tcgaactactctt 25  
 QY 941 TCAGAGACTACTCTT 954



Page 90

Db	59	tcataactctatt	73
	37	tcaatctgatt	51
oy			

[illegible]

WPI-29-24684/74.  
polymorphic bovine DNA markers used in genetic identification,  
Table 7, Page 396, 517pp; English.  
The sequence is that downstream of a bovine microsatellite sequen  
One hundred and thirty six clones of bovine DNA fragments of betwe  
200 and 500 bp were analysed by RFLP analysis using TaqI and HhaI  
One out of 50 clones cross-hybridised assuming independent  
distribution of microsatellites and MboI sites, the frequency of  
microsatellites was estimated as 1/1000.  
000". The sequence information for 200 such bovine microsatellit  
is summarised in the specification and indexed herein (see below).  
The sequences upstream and downstream of the microsatellite sequ  
amplification of the corresponding microsatellite using the program  
OPTIPRIM). The microsatellites may be used to identify individual  
for parentage testing, and in the genetic mapping of economic trait  
traits e.g. in cattle, to allow selective breeding of important

The sequences upstream and downstream of the microsatellite were used to generate the required PCR primers for use in vitro (see below). The microsatellites may be used to identify individuals for parentage testing, and in the genetic mapping of economic traits. The results involved the determination of economically important loci, which will allow selective breeding.

See also G13501-34437.  
Seq 631501-34437  
117 BP; 20 A; 25 C; 20 G; 52 T;

Query Match 1.0%; Score 14; DB 5; Length 117;  
Best Local Similarity 100.0%; Record No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0

6 CTGACATGGTCTCA 19

64 CTGAGATGTTACTG 77

## RESULT 71

Page 97

[illegible]

Sequence	118 BP	23 A	30 C	31 G	33 T
Query Match					
Score	1.38				
Best Local Similarity	100.0%				
Matches	15	Conservative	0	Mismatches	0
db	96 gccctccacgtatg	110			
Accession	U00001				
db	607 gccctccacgtatg	593			

Accession  
number  
T07133 standard; cDNA; 127 BP.  
T07133;  
TAC  
26-JUN-1996 (first entry)  
DP  
Vincular antimicrobial peptide coding sequence

[illegible]

US CLAIM 11; Page 23; 3/pp; English.



CC This sequence represents the cDNA encoding bovine epithelial lingual antimicrobial peptide (LAP). LAP is a member of the beta-defensin group of proteins that have been shown to have antimicrobial activity against Gram-positive and Gram-negative bacteria, and fungal pathogens. LAP is present at low levels in mammalian epithelia, with high expression levels being induced in response to epithelial injury/infection. This sequence can be used in a method of identifying and isolating genes that encode proteins that have antimicrobial activity cultured in the presence of a test substance. The levels of LAP mRNA are then measured to determine whether the substance is an up-regulator. LAP is used to treat microbial infections of the epithelium (or of those that are used to deposit microbes) e.g. in immunodeficient states (AIDS), cystic fibrosis, gum disease, burn, pneumonia, etc.

Query Match 1.38; Score 15; DB 18; Length 127;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 90 cctctcgcgcgcga 104  
Q7 581 cctctcgcgcgcga 595  
|||||

#### RESULT 74

ID Q62533 standard; DNA; 130 BP.  
AC 207114; 1996 (first entry)  
KW Antimicrobial peptide; epithelium; LAP; bovine; beta-defensin;  
KW Antimicrobial activity; Gram-positive bacteria; Gram-negative bacteria;  
KW Fungal pathogen; mammal; microbial infection; immunodeficiency; AIDS;  
KW Bovine; bovine; gum disease; burn; pneumonia; etc.  
Key signal\_peptide 84..89 Location/Qualifiers  
FT /tag= a  
FT /tag= b  
FT nat\_peptide 90..215 Pro-LAP

PP W0951297-A1.  
PP 15-JUN-1992; U06761.  
PP 24-MAY-1994; US-248016.  
PR (MAGA-) MAGALIN PHARM INC.  
DI Schoneveldt, R. S. J. /tag= b  
DI 15-JUN-1992; U06761.  
DR P-PSDB; R06895.  
PT New antimicrobial peptide from mammalian lingual epithelium - for treating bacterial or fungal infections, esp. of the epithelium, C. in gum disease, cystic fibrosis, burns, etc.  
CC This sequence represents the cDNA encoding the pro form of bovine epithelial lingual antimicrobial peptide (LAP). LAP is a member of the beta-defensin group of the defensin family of peptides. LAP has broad spectrum antimicrobial activity against Gram-positive and Gram-negative bacterial and fungal pathogens. LAP is present at low levels in mammalian epithelia, with high expression levels being induced in a response to epithelial injury/infection. This sequence can be used in a method of identifying endogenous up-regulators of LAP. In this method, epithelial cells are cultured in the presence of a test substance. The levels of LAP mRNA are then measured to determine whether the substance is an up-regulator. LAP is used to treat microbial infections of the epithelium (or of those that extend to deeper tissues) e.g. in immunodeficient states (AIDS), cystic fibrosis, gum disease, wounds, burns and pneumonia.

CC Burns and pneumonia.

Query Match 1.38; Score 14; DB 11; Length 130;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 tcgaactactctt 25

Q7 941 TCGAAGACTACTCTT 954

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Q7 Sequence 133 BP; 40 A; 27 C; 38 G; 28 T;  
Query Match 1.38; Score 15; DB 18; Length 133;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 96 cctctcgcgcgcga 110

Q7 581 cctctcgcgcgcga 595

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||



DE Human type I steroid 5-alpha reductase cDNA recombinant fragment.  
 KW probe; recombinant; inhibits prostatic hyperplasia; acute; hirsutism;  
 KW male pattern baldness; endometriosis; prostate cancer; testosterone;  
 KW dihydroxytestosterone; ss.  
 PN US542262-A.  
 PD 06-JUN-1995.  
 PF 30-APR-1990; 517661.  
 PR 30-APR-1990; US-517661.  
 PS Claim 1: A cDNA for the human type I steroid 5-alpha reductase  
 PA (TEXA) UNIV TEXAS SYSTEM.  
 PI Anderson S, Russell DW.  
 DR Steroid 5-alpha reductase nucleic acid segments and recombinant  
 PT cDNA. The cDNA segments are useful in e.g. analysis of normal  
 PT and abnormal sexual differentiation.  
 PS Claim 16: Column 61-62; 72pp; English.  
 CC 5-alpha reductase enzymes catalyze the conversion of testosterone to  
 CC dihydroxytestosterone. The human steroid 5-alpha reductase I (SRD5A-I)  
 CC cDNA has been cloned and sequenced. The cDNA encodes a protein of  
 CC this (C97482-097394) can be used in the preparation of probes for  
 CC for the large scale production of SRD5A or as probes for enzyme-encoding  
 CC sequences from alternative sources. The sequences are also useful in  
 CC the analysis of normal and abnormal sexual differentiation, benign  
 CC endometriosis, prostatic hyperplasia, male pattern baldness; acute; hirsutism and  
 CC endometriosis.  
 SQ Sequence 167 BP; 40 A; 33 C; 42 G; 52 T;

Query Match 1.38; Score 14; DB 17; Length 167;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 56 tctacatggcgat 69  
 Cp 419 tctacatggcgat 406  
 RESULT 79  
 ID T25902 standard; cDNA to mRNA; 175 BP.  
 AC W71324-1991 (first entry)  
 DE Probe (sub 1) to determine human individual genetic identity.  
 DE DNA probe; HeLa cell; HL-60; K-562; polymorphism;  
 OS Homo sapiens.  
 PN J6219530C-A.  
 PD 28-AUG-1987.  
 PF 19-FEB-1986; 034146.  
 PR 19-FEB-1986; JP-034146.  
 PS Claim 1: A cDNA for the human individual genetic identity.  
 PA (NICH) NICHIREI KK.

DR WPI: 87-280952.  
 PT DNA probe for assaying individual genetic identity - obtained from  
 PT HeLa cell DNA.  
 PS Claim 1: page 586; 6pp; Japanese.  
 CC The sequence is obtd. from HeLa cell DNA, HL-60 cell DNA, or K-562  
 CC cell DNA and is used for individual gene identity assay, by method  
 CC of hybridization with K562 cell DNA.  
 CC See also N70876 and N71324-28.  
 SQ Sequence 171 BP; 55 A; 31 C; 35 G; 50 T;  
 Query Match 1.38; Score 14; DB 14; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 52 tcagaactacttt 65  
 Cp 941 tcagaactacttt 954  
 RESULT 80  
 ID W71328 standard; DNA; 171 BP.  
 AC W71328-1991 (first entry)  
 DE Probe (sub 5) to determine human individual genetic identity.  
 DE DNA probe; HeLa cell; HL-60; K-562; polymorphism;  
 OS Homo sapiens.  
 PN J6219530C-A.  
 PD 28-AUG-1987.  
 PF 19-FEB-1986; 034146.  
 PR 19-FEB-1986; JP-034146.  
 PS Claim 1: A cDNA for the human individual genetic identity.  
 PA (NICH) NICHIREI KK.

Query Match 1.38; Score 14; DB 1; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 52 tcagaactacttt 65  
 Cp 941 tcagaactacttt 954  
 RESULT 83  
 ID G61339 standard; DNA; 216 BP.  
 AC G61339-1994 (first entry)  
 DE Human brain expressed sequence tag EST01339.  
 DE Gene transcription product; genetic markers; tagging; in vivo;  
 KW

DE Human type I steroid 5-alpha reductase cDNA recombinant fragment.  
 KW probe; recombinant; inhibits prostatic hyperplasia; acute; hirsutism;  
 KW male pattern baldness; endometriosis; prostate cancer; testosterone;  
 KW dihydroxytestosterone; ss.  
 PN US542262-A.  
 PD 06-JUN-1995.  
 PF 30-APR-1990; 517661.  
 PR 30-APR-1990; US-517661.  
 PS Claim 1: A cDNA for the human type I steroid 5-alpha reductase  
 PA (TEXA) UNIV TEXAS SYSTEM.  
 PI Anderson S, Russell DW.  
 DR Steroid 5-alpha reductase nucleic acid segments and recombinant  
 PT cDNA. The cDNA segments are useful in e.g. analysis of normal  
 PT and abnormal sexual differentiation.  
 PS Claim 16: Column 61-62; 72pp; English.  
 CC 5-alpha reductase enzymes catalyze the conversion of testosterone to  
 CC dihydroxytestosterone. The human steroid 5-alpha reductase I (SRD5A-I)  
 CC cDNA has been cloned and sequenced. The cDNA encodes a protein of  
 CC this (C97482-097394) can be used in the preparation of probes for  
 CC for the large scale production of SRD5A or as probes for enzyme-encoding  
 CC sequences from alternative sources. The sequences are also useful in  
 CC the analysis of normal and abnormal sexual differentiation, benign  
 CC endometriosis, prostatic hyperplasia, male pattern baldness; acute; hirsutism and  
 CC endometriosis.  
 SQ Sequence 167 BP; 40 A; 33 C; 42 G; 52 T;

Query Match 1.38; Score 14; DB 17; Length 167;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 56 tctacatggcgat 69  
 Cp 419 tctacatggcgat 406  
 RESULT 79  
 ID T25902 standard; cDNA to mRNA; 175 BP.  
 AC W71324-1991 (first entry)  
 DE Probe (sub 1) to determine human individual genetic identity.  
 DE DNA probe; HeLa cell; HL-60; K-562; polymorphism;  
 OS Homo sapiens.  
 PN J6219530C-A.  
 PD 28-AUG-1987.  
 PF 19-FEB-1986; 034146.  
 PR 19-FEB-1986; JP-034146.  
 PS Claim 1: A cDNA for the human individual genetic identity.  
 PA (NICH) NICHIREI KK.

DE Human type I steroid 5-alpha reductase cDNA recombinant fragment.  
 KW probe; recombinant; inhibits prostatic hyperplasia; acute; hirsutism;  
 KW male pattern baldness; endometriosis; prostate cancer; testosterone;  
 KW dihydroxytestosterone; ss.  
 PN US542262-A.  
 PD 06-JUN-1995.  
 PF 30-APR-1990; 517661.  
 PR 30-APR-1990; US-517661.  
 PS Claim 1: A cDNA for the human type I steroid 5-alpha reductase  
 PA (TEXA) UNIV TEXAS SYSTEM.  
 PI Anderson S, Russell DW.  
 DR Steroid 5-alpha reductase nucleic acid segments and recombinant  
 PT cDNA. The cDNA segments are useful in e.g. analysis of normal  
 PT and abnormal sexual differentiation.  
 PS Claim 16: Column 61-62; 72pp; English.  
 CC 5-alpha reductase enzymes catalyze the conversion of testosterone to  
 CC dihydroxytestosterone. The human steroid 5-alpha reductase I (SRD5A-I)  
 CC cDNA has been cloned and sequenced. The cDNA encodes a protein of  
 CC this (C97482-097394) can be used in the preparation of probes for  
 CC for the large scale production of SRD5A or as probes for enzyme-encoding  
 CC sequences from alternative sources. The sequences are also useful in  
 CC the analysis of normal and abnormal sexual differentiation, benign  
 CC endometriosis, prostatic hyperplasia, male pattern baldness; acute; hirsutism and  
 CC endometriosis.  
 SQ Sequence 167 BP; 40 A; 33 C; 42 G; 52 T;

Query Match 1.38; Score 14; DB 18; Length 179;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 8 ccagggtttgaca 21  
 Cp 697 ccagggtttgaca 684  
 RESULT 83  
 ID G61339 standard; DNA; 216 BP.  
 AC G61339-1994 (first entry)  
 DE Human brain expressed sequence tag EST01339.  
 DE Gene transcription product; genetic markers; tagging; in vivo;  
 KW







Tue Nov 17 08:55:26 1998

```

CC all the 3'-oriented cDNAs hydriase with specific mRNAs. Each library
CC different mRNAs in case to reflect accurately the relative abundance of
CC the particular tissue from which it was derived.
CC The appearance frequency of a given OS in a cDNA library can be
CC estimated by comparing the number of clones obtained from the OS
CC sequences as a means of diagnosing abnormal cell function of for
CC recognising different cell types.
SQ Sequence 257 BP; 88 A; 36 C; 14 G; 83 T;

Query Match 139; Score 14; DB 20; Length 257;
Best Local Similarity 100.0%; Pred.No. 3.47e+02; O: Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 104 aaatcagacgacg 117
Cp 385 aattcagacgacg 372

RESULT 88
ID Q59887 standard; cDNA: 260 BP.
AC Q59887;
UT 16-MAR-1994 (first entry)
KW Gene transcription product; genetic markers; tagging; in vivo;
RW Homo sapiens.
OS Homo sapiens.
PS 19-01-1994
PR 12-FEB-1993; U01394.
PP 12-FEB-1992; US-837195.
PA (USSH US DEPT HEALTH & HUMAN SERVICE.
PI NIH, Bethesda, Md., Venter CJ).
DA WPI: 93-272882/34.
DD WPI: 93-272882/34.
PT Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
PT Example 4; Page 263; 500pp; English.
PS The Expressed sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for human genes transcribed in vivo. They can be used to facilitate
CC identification of genes expressed in various tissues and mapped
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prepog. of antisense sequences, probes and constructs.
CC This information will be useful for identifying the location of the
CC coding-region prediction program CRM. See also Q59041-Q61440.
SQ Sequence 260 BP; 63 A; 34 C; 14 G; 111 T;

Query Match 139; Score 14; DB 8; Length 260;
Best Local Similarity 93.3%; Pred.No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Page 107

Tue Nov 17 08:55:26 1998

Q# 499 CTCACAGCCCCC 486

NR007 90  
IDT T31602 standard; cDNA to mRNA; 293 BP.  
DE T31602;  
CC  
DT 02-SEP-1996 (first entry)  
DE Human gene signature HMG850548.  
KW Human cloning; mapping; non-biased library; diagnosis; detection;  
KS cell typing; abnormal cell function; ss.  
OS Homo sapiens.  
PN OJ293121.  
PF 03-JUN-1995.  
PP 11-NOV-1994; J01916.  
PR 12-NOV-1993; JP-355504.  
PA (MATSU) MATSUHARA K.  
PI Matsuhara K., Hiroko Okubo Ki,  
DR WPI: 95-265931/27.  
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT identifying gene signatures in 3'-directed human cDNA library that  
PT identifies relative abundance of corresp. mRNA in specific human  
PT tissues  
PS Claim 1. Page 1409; 2245pp; Japanese.  
CC double-stranded DNA (c) to complementary strand or the corresp.  
CC given in T31601-726817 and which is able to hybridize to part of  
CC human genomic DNA. cDNA or mRNA is claimed. The GS (Gene Signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC at the 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNA hybridise with specific mRNAs. Each library  
CC contains different mRNAs in the particular tissue. The relative abundance of  
CC different mRNAs in the particular tissue can be estimated.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing animal cell function or for  
CC screening of genes.  
SQ Sequence : 293 BP; 82 A; 82 C; 43 G; 56 G; 98 T;

Query Match 13; Score 14; DB 20; Length 293;  
Local Similarity 10%; Pval: 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

D# 96 gaattctgttaca 109  
GT 663 GAATTTGTTGTA 676

Page 108

Thu Nov 17 08:55:26 1998  
UG-08-887-977-9, P88

[illegible]







Page 116

Page 115

[illegible]

W09311237-A.  
10-JUN-1993.  
04-DEC-1992. G02251.  
04-DEC-1992. G02251.  
PFA (SHEP) WILCOX FOUND. LTD.  
Crown JS, Lewis AP, Waldman H, Walsh LJ  
WPI: 93-197056/P24.  
P-PEDM: R38159.  
CDR-grated humanised chimeric T cell antibodies - inhibit T cell proliferation, interfering with cell mediated diseases e.g. rheumatoid arthritis, autoimmune-diseases etc.  
Claim 9; Page 27; 49pp; English.  
The CDRA are of the anti-human T cell antibody YH 655(516). YH 655(516) is a rat IgG2b monoclonal antibody which binds to resting and activated T cells, PCR reactions were carried out to humanise the rat Mab YH 655. The human variable domain frameworks of the MOI humanisation process. Humanised YH 655 heavy and light chains were cloned into a eukaryotic expression vector under human cytomegalovirus promoters and expressed transiently in COS cells at 200mg/ml. The humanised YH 655 was shown by FACS analysis to bind the activated T cell binding site NP14. The binding to NP14 cells was quantified by flow cytometry. Sequence 130 BP. Length 84 CT. 81 T.

Page 116



Query Match 1.38; Score 14; DB 7; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 251 agctggagctgctt 264  
 Cp 1009 ACCCGGAGCTG 296

RESULT 101  
 ID T27968 standard; DNA: 340 BP.  
 AC Q77282;  
 DT 23-SEP-1994 (first entry)  
 DE Human genome fragment (Preferred).  
 KW Hepatitis C virus; genotype; polymerase chain reaction; amplification;  
 KW PCR; primer; probe; antibody; infection; ss.  
 OS Homo sapiens.  
 PN W091472-A.  
 PP 11-JUN-1995; J01916.  
 PR 12-NOV-1993; JP-355504.  
 PA (MATSU) MATSUBARA K.  
 PA (ORUB) ORUBO K.  
 DA WP1: 95-206931/27; Subo K.  
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
 PT for diagnosis of abnormal cell function, by preparing cDNA that  
 PT reflects relative abundance of corresp. mRNA in specific human  
 cell typing; abnormal cell function; ss.  
 CC Claim 1. Page 931: 2245pp; Japanese.  
 CC A single-stranded DNA (or its complementary strand or the corresp.  
 CC double-stranded DNA) which comprises one of the 7837 "Gs" sequences  
 CC given in T19001-72817 and which is able to hybridize to part of  
 CC sequences were obtained from 3'-directed cDNA libraries prepared  
 CC from various human tissues; synthesis of cDNA was initiated from the  
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence unique to a particular mRNA species, ribost  
 CC is constructed so as to reflect accurately the relative abundance of  
 CC different mRNAs in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC sequenced as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types. 62 C; 41 G; 120 T;  
 CC Sequence 335 BP; 108 A; 62 C; 41 G; 120 T;

Query Match 1.38; Score 14; DB 19; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

cell typing; abnormal cell function; ss.  
 OS Homo sapiens.  
 PN W091472-A.  
 PP 11-JUN-1995; J01916.  
 PR 12-NOV-1993; JP-355504.  
 PA (MATSU) MATSUBARA K.  
 PA (ORUB) ORUBO K.  
 DA WP1: 95-206931/27.  
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
 PT for diagnosis of abnormal cell function, by preparing cDNA that  
 PT reflects relative abundance of corresp. mRNA in specific human  
 cell typing; abnormal cell function; ss.  
 CC Claim 1. Page 2095: 2245pp; Japanese.  
 CC A single-stranded DNA (or its complementary strand or the corresp.  
 CC double-stranded DNA) which comprises one of the 7837 "Gs" sequences  
 CC given in T19001-72817 and which is able to hybridize to part of  
 CC sequences were obtained from 3'-directed cDNA libraries prepared  
 CC from various human tissues; synthesis of cDNA was initiated from the  
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence unique to a particular mRNA species, ribost  
 CC is constructed so as to reflect accurately the relative abundance of  
 CC different mRNAs in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC sequenced as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types. 61 C; 62 G; 114 T;  
 CC Sequence 340 BP; 17 A; 61 C; 62 G; 114 T;

Query Match 1.38; Score 14; DB 23; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 267 cttttgtttttt 280  
 Qy 631 CTTTGTGTTT 644

RESULT 104  
 ID Q31044 standard; DNA: 340 BP.  
 AC Q31044;  
 DT 24-MAR-1993 (first entry)  
 DE HCV-1 genotype GII NSS region sequence na89h.  
 KW Hepatitis C virus; non-A, non-B hepatitis; NSS region; ss.  
 PN W09219743-A.  
 PP 12-NOV-1992.  
 PR 08-MAY-1992; 004036.

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 32 ttgtgttttttt 45  
 Qy 194 TTGCTTTTATTA 197

RESULT 102  
 ID Q77282 standard; DNA: 337 BP.  
 AC Q77282;  
 DT 23-SEP-1994 (first entry)  
 DE Human genome fragment (Preferred).  
 KW Hepatitis C virus; genotype; polymerase chain reaction; amplification;  
 KW PCR; primer; probe; antibody; infection; ss.  
 OS Homo sapiens.  
 PN W091472-A.  
 PP 11-JUN-1995; J01916.  
 PR 12-NOV-1993; JP-355504.  
 PA (MATSU) MATSUBARA K.  
 PA (ORUB) ORUBO K.  
 DA WP1: 95-206931/27; Subo K.  
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
 PT for diagnosis of abnormal cell function, by preparing cDNA that  
 PT reflects relative abundance of corresp. mRNA in specific human  
 cell typing; abnormal cell function; ss.  
 CC Claim 1. Page 931: 2245pp; Japanese.  
 CC A single-stranded DNA (or its complementary strand or the corresp.  
 CC double-stranded DNA) which comprises one of the 7837 "Gs" sequences  
 CC given in T19001-72817 and which is able to hybridize to part of  
 CC sequences were obtained from 3'-directed cDNA libraries prepared  
 CC from various human tissues; synthesis of cDNA was initiated from the  
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence unique to a particular mRNA species, ribost  
 CC is constructed so as to reflect accurately the relative abundance of  
 CC different mRNAs in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC sequenced as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types. 88 A; 83 C; 55 G; 111 T;  
 CC Sequence 337 BP; 88 A; 83 C; 55 G; 111 T;

Query Match 1.38; Score 14; DB 11; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 actggtgtaattga 52  
 Cp 58 ACTGAGTATATGA 45

RESULT 103  
 ID T26475 standard; cDNA to mRNA. 340 BP.  
 AC T26475;  
 DT 23-OCT-1996 (first entry)  
 DE Human gene signature HMG080717.  
 KW Hepatitis C virus; genotype; polymerase chain reaction; amplification;  
 KW PCR; primer; probe; antibody; infection; ss.  
 OS Homo sapiens.  
 PN W091472-A.  
 PP 11-JUN-1995; J01916.  
 PR 12-NOV-1993; JP-355504.  
 PA (MATSU) MATSUBARA K.  
 PA (ORUB) ORUBO K.  
 DA WP1: 95-206931/27; Subo K.  
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
 PT for diagnosis of abnormal cell function, by preparing cDNA that  
 PT reflects relative abundance of corresp. mRNA in specific human  
 cell typing; abnormal cell function; ss.  
 CC Claim 1. Page 931: 2245pp; Japanese.  
 CC A single-stranded DNA (or its complementary strand or the corresp.  
 CC double-stranded DNA) which comprises one of the 7837 "Gs" sequences  
 CC given in T19001-72817 and which is able to hybridize to part of  
 CC sequences were obtained from 3'-directed cDNA libraries prepared  
 CC from various human tissues; synthesis of cDNA was initiated from the  
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence unique to a particular mRNA species, ribost  
 CC is constructed so as to reflect accurately the relative abundance of  
 CC different mRNAs in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC sequenced as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types. 88 A; 83 C; 55 G; 111 T;  
 CC Sequence 337 BP; 88 A; 83 C; 55 G; 111 T;

08-MAY-1991; US-697326.  
 PR (CHIR) CHIRON CORP.  
 PI Beall E, Cha T, Irvine B, Kolberg J, Urdea MS;  
 PI 1991; 10: 101-104.  
 PT Comparing a non-hepatitis C virus-1 nucleotide sequence  
 PT - related to HCV-1, useful for treating and detecting HCV-1  
 PT infections and as a vaccine  
 PS Claim 4: Page 62; 18pp; English.  
 CC The invention relates to a nucleotide  
 CC sequence within the NSS region of HCV-1 of genotype GII. The sequence  
 CC is a preferred example of an oligonucleotide for use as a probe in  
 CC hybridisation assays, as a primer for synthesis of HCV genotype-  
 CC specific nucleic acid, as a binding partner for separating HCV  
 CC expression of HCV genes. Polypeptides encoded by oligonucleotides of  
 CC the invention (no sequences given in the specification) are useful  
 CC in vaccines against HCV and to produce antibodies to detect the  
 CC virus.  
 CC Sequence 340 BP; 75 A; 97 C; 104 G; 64 T;

Query Match 1.38; Score 15; DB 5; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 ttctgtctatctgc 299  
 Cp 1076 TTCTGTCTATCTGC 1062

RESULT 105  
 ID T27968 standard; DNA: 340 BP.  
 AC T27968;  
 DT 23-SEP-1994 (first entry)  
 DE Human genome fragment (Preferred).  
 KW Hepatitis C virus; genotype; polymerase chain reaction; amplification;  
 KW PCR; primer; probe; antibody; infection; ss.  
 OS Homo sapiens.  
 PN W091472-A.  
 PP 11-JUN-1995; J01916.  
 PR 12-NOV-1993; JP-355504.  
 PA (MATSU) MATSUBARA K.  
 PA (ORUB) ORUBO K.  
 DA WP1: 95-206931/27; Subo K.  
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
 PT for diagnosis of abnormal cell function, by preparing cDNA that  
 PT reflects relative abundance of corresp. mRNA in specific human  
 cell typing; abnormal cell function; ss.  
 CC Claim 1. Page 931: 2245pp; Japanese.  
 CC A single-stranded DNA (or its complementary strand or the corresp.  
 CC double-stranded DNA) which comprises one of the 7837 "Gs" sequences  
 CC given in T19001-72817 and which is able to hybridize to part of  
 CC sequences were obtained from 3'-directed cDNA libraries prepared  
 CC from various human tissues; synthesis of cDNA was initiated from the  
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence unique to a particular mRNA species, ribost  
 CC is constructed so as to reflect accurately the relative abundance of  
 CC different mRNAs in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC sequenced as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types. 88 A; 83 C; 55 G; 111 T;  
 CC Sequence 337 BP; 88 A; 83 C; 55 G; 111 T;

Query Match 1.38; Score 14; DB 23; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 267 cttttgtttttt 280  
 Qy 631 CTTTGTGTTT 644

RESULT 104  
 ID Q31044 standard; DNA: 340 BP.  
 AC Q31044;  
 DT 24-MAR-1993 (first entry)  
 DE HCV-1 genotype GII NSS region sequence na89h.  
 KW Hepatitis C virus; non-A, non-B hepatitis; NSS region; ss.  
 PN W09219743-A.  
 PP 12-NOV-1992.  
 PR 08-MAY-1992; 004036.



PT Hepatitis C virus poly(nucleic acid unique to unidentified sub-type  
PT - used to develop probes and primers for new sub-types and vaccines  
PS Claim 6: Fig 3: 150pp; English.  
CC The sequences 727937-727989 represent novel sequences isolated from  
CC human liver tumours of the HCV type 1a, 2a, 4, 2b, 1-12,  
CC 3a, 4a-m, 7a-c or types 9, 10 or 11. The sequences correspond to the 5',  
CC untranslated region (UR), the Core/2a, NS4 or NS5B regions of the genome.  
CC This sequence represents nucleotides 7932-8271 from the HCV type 1a  
CC The new HCV types were isolated from patients with chronic HCV from the  
CC Benelux countries, France, Cameroon and Vietnam, because of their  
CC aberrant reactivities. The RNA was extracted, cDNA synthesised and PCR  
CC amplified, cloned and genotyped. The 5'UR, Core/2a and NS5B regions were  
CC sequenced and compared with the known sequences. The new  
CC viruses into (sub)types based on comparison with known sequences.  
CC The sequences were used to generate the peptides R9624-R96524. The  
CC sequences can also be used to synthesise probes and primers for the  
CC detection of HCV in a sample. The polypeptides can be used to detect  
CC HCV infection and/or to develop vaccines for the prevention of infections.  
SQ Sequence 340 BP: 77 A: 94 C: 92 G: 63 T: 1

Query Match 1.38; Score 15; DB 25; Length 340;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 ttgtcttctctctg 299  
CP 1076 TTTCTCTTCTCTCTG 1062

RESULT 106  
ID 724517 standard: cDNA to mRNA, 340 BP.  
AC 724517-1996 (first entry)  
DE Human gene signature HMG0506560.  
KW Gene signature: messenger RNA: mRNA: relative abundance; frequency;  
KW human; cloning; mapping; non-biased library; diagnosis; detection;  
KW cell typing; abnormal cell function; ss.  
OS Cell typing; abnormal cell function; ss.  
PN W09314772-A1.  
PD 01-JUN-1995.  
PF 11-NOV-1994; J01915.  
PR 11-NOV-1994; J01915.  
PA (MAGN.) MATSUURA K.  
PA (OKUBO.) OKUBO K.  
PI Matsubara K, Okubo K;  
DR W92: 92-06993/27; signatures in 3'-directed human cDNA library - e.g.  
DR identifying gene transcripts in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human  
PS Claim 1: Page 1629: 2245pp; Japanese.  
CC A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "Gs" sequences  
CC given in 719001-728937 and which is able to hybridise to part of  
CC human genomic DNA; cDNA or mRNA from a mammalian cell (Gene Signature)  
CC 3'-end of mRNA by using poly(7) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the translated sequences of the corresponding mRNA species are  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC recombinant different segments of diagnosing abnormal cell function or for  
SQ Sequence 340 BP: 98 A: 57e+42 C: 73 G: 109 T:

Query Match Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 187 gaattgtttctga 200  
CP 1043 GAAATGTTTCTGA 1030

RESULT 107  
ID 000487 standard: cDNA; 342 BP.  
AC 000487-1994 (first entry)  
DE Human brain expressed sequence tag EST02491.  
KW Gene transcription product; genetic markers; tagging; in vivo;  
KW transcription; mapping; locations; chromosomes; chromosomal; ss.  
OS Transcription; mapping; locations; chromosomes; chromosomal; ss.  
PN W09316178-A1.  
PD 19-AUG-1993.  
PF 12-FEB-1993; U01294.  
PR 12-FEB-1993; U01294.  
PA (MAGN.) MATSUURA K.  
PA (OKUBO.) OKUBO K.  
PI Matsubara K, Okubo K;  
DR W92: 93-37282/34; signatures in 3'-directed human cDNA library - e.g.  
DR identifying gene transcripts in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that

CC Tagging of most human genes, for mapping locations of expressed genes  
CC on chromosomes, for individual or forensic identification, for mapping  
CC locations of disease-associated genes, for identification of tissue  
CC expression or of protein or polypeptide sequences, probes and constructs.  
CC EST02491 has been deposited with the European Molecular Biology  
CC coding-region prediction program CDM. See also Q39041-Q81440.  
SQ Sequence 342 BP: 98 A: 97 C: 78 G: 65 T: 1

Query Match 1.38; Score 14; DB 8; Length 342;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 acagcccccacaca 237  
CP 496 ACAGCCCCCACA 483

RESULT 108  
ID 769400 standard: cDNA; 350 BP.  
AC 769400-1996 (first entry)  
DE Murine metastatic nucleic acid sequence.  
DE Mouse; murine; tumour; cancer; metastatic sequence; detection;  
KW hyperplasia; screening; ss.  
OS Mus musculus.  
PN W09718454-A2.  
PD 15-NOV-1995; U18567.  
PR 10-JAN-1996; US-594031.  
PA (THOM.) THOMPSON T.  
PI (THOM.) THOMPSON T.  
DR W92: 92-28939/26; signatures in 3'-directed human cDNA library - e.g.  
DR identifying tumour metastatic sequences - by introducing transfected  
PT cells into host mammal and analyzing primary and metastatic  
PT sequences by differential display PCR  
PT sequences by differential display PCR  
CC Mouse embryonal sinus (UES) tissues were isolated from 17 day old  
CC mouse embryos. The UES cells were infected with retroviruses,  
CC cultured and implanted under the renal capsule. RNA was isolated  
CC from a site outside the renal capsule. RNA was isolated  
CC from primary tumours and metastases, reverse transcribed and  
CC subjected to differential display PCR. The sequences were analysed  
CC to obtain metastatic sequences, e.g. the present sequence. The  
CC sequences were used to screen a cDNA library for metastatic sequences,  
CC to metastasis, or treat malignant or non-malignant disorders, e.g.  
CC hyperplasia, dysplasia and hypertrophy. The metastatic sequence can  
CC be used to screen a biological sample for metastasis, and it or its

CC expression product may also be used to treat a metastatic disorder.  
SQ Sequence 350 BP: 114 A: 80 C: 74 G: 82 T: 1

Query Match Similarity 100.0%; Pred. No. 8.37e+01;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 331 catctgcctgtctgt 345  
CP 471 CATCTGCTGTCTGT 485

RESULT 109  
ID 707135 standard: cDNA; 350 BP.  
AC 707135-1996 (first entry)  
DE Prepro-LAP coding sequence.  
DE Lingual antimicrobial peptide; epithelium; LAP; bovine; beta-defensin;  
KW antimicrobial activity; Gram-positive bacteria; Gram-negative bacteria;  
KW cystic fibrosis; gum disease; burn; pneumonia; ss.  
OS Bos taurus.  
PN Key Location/Qualifiers  
PF 24..218  
PR 24..218  
PA (MAGN.) MATSUURA K.  
PA (OKUBO.) OKUBO K.  
PI Matsubara K, Okubo K;  
DR W92: 92-06993/27; signatures in 3'-directed human cDNA library - e.g.  
DR identifying gene transcripts in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that

Query Match 1.38; Score 15; DB 25; Length 340;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 187 gaattgtttctga 200  
CP 1043 GAAATGTTTCTGA 1030

RESULT 107  
ID 000487 standard: cDNA; 342 BP.  
AC 000487-1994 (first entry)  
DE Human brain expressed sequence tag EST02491.  
KW Gene transcription product; genetic markers; tagging; in vivo;  
KW transcription; mapping; locations; chromosomes; chromosomal; ss.  
OS Transcription; mapping; locations; chromosomes; chromosomal; ss.  
PN W09316178-A1.  
PD 19-AUG-1993.  
PF 12-FEB-1993; U01294.  
PR 12-FEB-1993; U01294.  
PA (MAGN.) MATSUURA K.  
PA (OKUBO.) OKUBO K.  
PI Matsubara K, Okubo K;  
DR W92: 93-37282/34; signatures in 3'-directed human cDNA library - e.g.  
DR identifying gene transcripts in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that



Query Match 1.34; Score 15; DB 10; Length 354;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 328 ttcattacttcagt 342  
|||||  
QY 44 TCAATATTCAGT 58

RESULT 111  
ID 122895; standard: cDNA to mRNA, 358 BP.  
DT 29-AUG-1996 (first entry)  
DE Human gene signature HMGSO4614.  
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
KW Human cloning; mapping; non-biased library; diagnosis; detection;  
KW Cell type; abnormal cell function; ss.  
OS Homo sapiens.  
PB W0514772-A1.  
PD 01-JUN-1995. 303936  
PR 12-NOV-1984. JP-355504.  
PA (MATSU) MATSUURA K.  
PA (MATSU) OKUBO K.  
DI Matsura, K. Okubo K.  
PI Scott NS, Thomas MR.  
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT tissue.  
PA Claim 1: Page 1236; 2245pp; Japanese.  
CC A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
CC given in T19001-T76837 and which is able to hybridise to part of  
CC human genomic DNA; cDNA; mRNA; or a derivative thereof, is used to  
CC sequence genomic DNA; cDNA; mRNA; or a derivative thereof, to  
CC identify human genes; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(7) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all mRNAs in a particular tissue or cell type can be identified.  
CC This is constructed so as to reflect accurately the abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC sequences) and used for identifying abnormal cell function or for  
CC recognising different cell types.  
SQ Sequence 358 BP; 67 A; 71 C; 88 G; 114 T;

Query Match 1.34; Score 14; DB 20; Length 358;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;

CC response to epithelial injury/infection. This sequence can be used in a  
CC method of identifying endogenous up-regulators of LAP. In this method,  
CC levels of LAP mRNA are then measured to determine whether the substance  
CC is an up-regulator. LAP is used to treat microbial infections of the  
CC epithelium (or of those that extend to deeper tissues) e.g. in  
CC immunodeficient states (AIDS), cystic fibrosis, gum disease, wounds,  
CC and other conditions.  
SQ Sequence 350 BP; 120 A; 72 C; 85 G; 73 T;

Query Match 1.34; Score 15; DB 18; Length 350;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 179 ctgtctcgagccca 193  
|||||  
QY 581 CTGCTCGAGCCCA 595

RESULT 110  
ID Q55233; standard: DNA; 354 BP.  
DT 14-JUL-1994 (first entry)  
DE Grapevine ribosomeal clone 176A.  
KW Ribosome; grapevine; vitis; IGS region; rDNA; polymorphism;  
KW Grape cultivar; probe; primer; detection; ds.  
OS Vitis vinifera. Location/Qualifiers  
FH Key primer\_bind 202..224  
FT primer\_bind /\*tag= a  
FT primer\_bind /\*tag= b  
PB W09401580-A.  
PD 20-JAN-1994.  
PR 30-JUN-1993; AU0120.  
PA (GSR) COMMONWEALTH SCI & IND RES ORG.  
PI Scott NS, Thomas MR.  
DR WPI: 94-035083/04.  
PT Novel ribosome DNA probe sequences - for the accurate identification  
PT of grapevine ribosomeal clones.  
PA Claim 26: Fig 5c; 55pp; English.  
CC The sequences given in Q55231-50 are derived from ribosomes of the  
CC grapevine genus Vitis. These sequences represent the IGS region of  
CC the ribosomeal rDNA repeat and contain polymorphisms. These  
CC polymorphisms are used to identify the ribosomeal clones of  
CC different grape cultivars. In these clones contain simple repeat  
CC sequences and were identified in a genomic library of grapevine DNA  
CC using simple di-, tri- or tetra- nucleotide repeats such as (AT)8.  
CC (GT)10, (CT)10 and such like as probes. 70 G; 92 T;  
SQ Sequence 354 BP; 117 A; 75 C; 70 G; 92 T;

QY 342 CACTATGCGATCA 355

RESULT 112  
ID Q99284; standard: DNA; 360 BP.  
DT 08-APR-1996 (first entry)  
DE Rat insulin-II gene extended 5'-flanking region.  
KW Rat; insulin; tissue-specific gene expression; Promoter;  
KW Human cloning; mapping; non-biased library; diagnosis; detection;  
KW Gene transfer; transcription; control; cis-acting sequence; ss.  
OS Rattus rattus.  
PB Key Location/Qualifiers  
FT Key Location/Qualifiers  
FT Data\_signal 331..337  
FT Data\_signal /\*tag= a  
PB US5486462-A.  
PD 23-JAN-1996.  
PR 23-NOV-1984; 674225.  
PR 18-MAY-1988; US-196781.  
PR 22-MAR-1989; US-327366.  
PR 30-SEP-1991; US-768345.  
PR 18-NOV-1991; US-700512.  
PR (RSCG) 1991; US-700512.  
PI Boulet AM, Edmund T, Rutter WJ, Wlaker MD.  
WT: 96-097074/10.  
PT Tissue-specific transcription-enhancing DNA segment - isolated from  
PT the 5'-flanking region of insulin, chymotrypsin or amylase gene, used in  
PT recombinant vector to enhance expression of heterologous protein in  
PT host cell.  
PA Disclosure: Fig 4; 16pp; English.  
CC The sequence represents an extended 5'-flanking region of a  
CC gene. The sequence is a partial, and includes a promoter.  
CC Sequences of this type may be identified using a vector with a  
CC reporter gene, and used in tissue-specific gene expression in  
CC insulin-producing mammalian cells, e.g. for heterologous insulin  
CC production. The sequence is a partial, and includes a promoter.  
CC 5'-flanking regions (from -1 to -300 bases upstream of the  
CC transcription initiation site) contain cis-acting sequences  
CC responsive to indigenous trans-acting substances in the host cell,  
CC giving enhanced heterologous protein expression.  
SQ Sequence 360 BP; 90 A; 96 C; 92 G; 78 T;

Query Match 1.34; Score 14; DB 17; Length 360;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;

DB 36 gctctcgagccat 49  
|||||  
QY 152 GCTCTCGAGCAT 165

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 229 cttgtcttttata 242  
|||||  
QY 183 CTTGCTTTTATA 196

RESULT 112  
ID Q41128; standard: DNA; 359 BP.  
DT 20-AUG-1993 (first entry)  
DE Human small lung cell carcinoma clone SCLL-A.  
KW SCLC; calcium channel; ss.  
OS Homo sapiens. Location/Qualifiers  
FH Key misc\_feature 12..359  
FT misc\_feature /\*tag= a  
FT misc\_feature /\*tag= b  
PB W09100469-A.  
PD 29-APR-1993.  
PR 21-OCT-1992; U09109.  
PR 23-OCT-1991; US-781421.  
PR 18-MAY-1988; US-196781.  
PI Barry RL, Probstner SC.  
WT: 93-152621/18.  
PT Recombinant alpha-1 subunit of voltage-sensitive calcium channel  
PT protein - for diagnosing small cell lung carcinoma, and treating  
PT syndromes etc. and for diagnosing and treating Lambert-Eaton  
PT syndrome.  
PA Claim 9; page 28; 49pp; English.  
CC The SCLL-A clone represents one of three types of putative calcium  
CC channel found by PCR amplification of a cDNA template prepared from  
CC human small lung cell carcinoma tissue. The sequence is a partial,  
CC perfectly or almost perfectly conserved between the rabbit skeletal  
CC muscle and cardiac calcium channels. SCLL-B and SCLL-C clones are  
CC highly homologous to each other and to the skeletal muscle calcium  
CC channel. The degree of homology to the cardiac channel is more  
CC than to the skeletal muscle channel. The sequence is a partial,  
CC recently characterised rabbit brain calcium channel, and is thought  
CC to represent a human neuronal calcium channel. The specification  
CC includes an amino acid sequence (see R34549). The C-terminal region  
CC of Q41128 is directly translated from Q41128.  
SQ Sequence 359 BP; 75 A; 92 C; 81 G; 111 T;

Query Match 1.34; Score 14; DB 7; Length 359;  
Best Local Similarity 100.0%; Pred. No. 3.47e+01;

DB 50 cactatgacatca 63  
|||||  
QY 50 CACTATGACATCA 63



ID	Q60260 standard; DNA; 387 BP.
AC	Q60260;
DT	16-MAR-1994 (first entry)
DE	Human brain Expressed sequence Tag ESTO2257.
DD	Human brain Expressed Sequence Tag database; tagging; in vivo;
DS	transcription; mapping; localization; chromosomal; ss.
KW	Homo sapiens.
PX	W09316178-A.
PN	13-NOV-1993;
PD	12-FEB-1992; US-837195.
PP	(US83) : US DEPT HEALTH & HUMAN SERVICE.
PR	Adams MD, Moreno RP, Venter CJ.
PI	Enriched oligonucleotides and corresp. sequences - used as
PF	markers for human genes transcribed in-vivo, facilitate tagging
PT	of most human genes
PS	library #; Page 306; 500pp; English.
CC	Example 4; Part of a large set of ESTs which can be used to facilitate
CC	tagging of most human genes transcribed in vivo. They can be used to facilitate
CC	localization or identification of genomic identification, for mapping
CC	location and/or expression of genes. The library contains probes and constructs
CC	type, and for prep. of antisense sequences, probes and constructs
CC	ESTO2257 has a "marginal" coding probability as evaluated using the
CC	coding-region prediction program CRM. See also Q5041-Q5140.
CC	Sequence 387 BP; 99 A; 85 C; 84 G; 118 T.
SQ	Query Match 1.3%; Score 14; DB 8; Length 387; Best Local Similarity 100.0%; Pred. No. 3,47e+02; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps
Df	36 attcttcactcgag 49
Cp	24 AAATCTACTGAG 11       
RE	RESULT 116
ID	YD24838 standard; DNA; 396 BP.
AC	YD24838;
DT	16-JUN-1988 (first entry)
DE	E. Pylori OMP hp301086orf2.
DS	Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW	Identification; binding compound; bacteria; life cycle; activator;
KN	Helicobacter pylori ulcer disease, chronic gastritis; diagnosis; ds.
PH	Location/Qualifiers
FT	1..396
FT	/contig 8
FT	"an open codon given"

```

DE VTR locus D18g17 flanking regions.
OS variable number tandem repeat; oligonucleotide probe.
FH Homo sapiens.
Key repeat_region      Location/Qualifiers
FT   1            406       /tag= a
FT   2            234..290  /note='sec comment'
FT   3            231..347  /tag= b
FT   4            231..347  /tag= c
FT   5            231..347  /number= 1
FT   6            231..347  /number= 2
FT   7            231..347  /number= 3
FT   8            405..461  /tag= d
FT   9            405..461  /tag= e
FT   10           405..461  /number= 4
FN W09110748.A.
PD 25-JUN-1991.
PD 09-JUN-1991. DO019660.
PR 19910601.
PA (COLS ) COLLABORATIVE RES INC.
KE Keith TP, Mao J, Rose SD;
WR 91-238036/32. from specific human genomic loci - used for
FR probes in the identification of individuals in forensic applications.
PE Disclosure: Page 23: 49pp; English.
CC The sequence was detected by probe CR1-L159-2 and comprises the 5'
CC flanking region and the first 4 tandem repeat units of a section of
CC the gene. This probe has been used to identify individuals in the
CC region and the last 4 tandem repeat units are given in Q13169.
CC Additional unsequenced repeats are present between these two
CC sequences. See also Q12805-Q12913 and Q13613-Q13617.
SQ Sequence 404 BP; 100 A; 143 C; 71 G; 90 T;
Query Match          1.3% Score 15; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps
DB 74 gcagctgcgatcgtt 88
CP |-----|
CP 337 GCAGCTGCATTCCT 313

```







[illegible]

```

UT-08-087-977-9.rms

      intron      1..72      /*tag= b
      /*number= 8
      /*note= "fragment of intron 8"
      exon        73..144    /*tag= C
      /*number= 9

W0626268-A1.
29-ADG-1996.  W04320.
21-FEB-1995.  US-351743.
(GEN2 ) GENZYME TRANGENICS CORP.
PI Cole ES, Dittullo P, Meade H.
DR W12; 96-40256/40.
CC acquired or inherited AT3 deficiency, with faster clearance rate
CC than plasma AT3 10. 37pp; Enghish.
CC Disclosure: Fig 10. 37pp; Enghish.
CC T5929-T5985 represent fragments of the goat beta-casein gene, and
CC T5929-T5985 represent fragments of the goat beta-casein gene, and
CC used to produce the transgenic antithrombin III (AT3) gene action.
CC AT3 is a serine protease inhibitor, which inhibits thrombin and the
CC activated forms of factors V, VII, IX, XI, and XII. The transgenic AT3
CC of the invention includes a monosaccharide composition containing
CC transgenic human AT3 cDNA (cloned as product of primary gland specific
CC into the goat beta-casein gene. The final 14.95 kb vector was
CC microinjected into goat embryos, which when developed produced the AT3
CC in their milk. The AT3 lacks O-linked glycosylation, and the major
CC except the Asn155 residue. The AT3 is used in the same way as
CC oligomannose, and hybrid residues. The AT3 is used in the same way as
CC plasma AT3, for the treatment of hereditary or acquired AT3 deficiency.
CC The AT3, however, has a faster clearance rate than plasma AT3 and may
CC without the variability, immunogenicity and viral contamination problems
CC associated with plasma-derived material.
CC Sequence 494 BP: 176 A: 62 C: 62 G: 193 T:
CC
CC Query Match 1.39; Score 15; DB 28; Length 494;
CC Best Local Similarity 100.00; Prod No. 837e+01;
CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
Db 282 aaatcatgaact 296
Cc 669 AATATCATGAACAT 655

RESULT 135
INFO 01455
AC Q14952;

```

```

DE P. putida LMG 2232 16S-23S rRNA spacer region.
PE Probe detection: Identification: microorganisms; amplify:
KW 16S-23S rRNA spacer region; respiratory tract; universal;
KW Pseudomonas sp.
OS Pseudomonas sp.
OS W09600238.A1.
PD PD
PD 04-JUN-1996. 807452
PD 24-JUN-1994. E2-8770106
PD 24-JUN-1994. E2-8770106
PR PF-APR-1993; EF-870032.
PR (INNO-) INNOGENETICS NV.
PA Janssens G, Rossau R, Van Heuverswyn H:
PA Novel hybridization assay for the detection of subcutanea - esp
PT Mycobacterium, comprises amplifying the 16S-23S rRNA spacer
PT Claim 2, Fig 40; 218pp; English.
PS The sequences given in 111791-857 represent the 16S-23S rRNA spacer
PS region.
PS The sequences given in 111791-857 represent the 16S-23S rRNA spacer
PS region.
CC used in the method of the invention for the detection and identification
CC of at least one or more microorganisms. The method comprises amplifying
CC the 16S-23S rRNA spacer region and hybridizing one or more of the probes
CC given in 114011-7 to the amplified sequence. These probes were
CC given in 114011-7 to the amplified sequence. These probes were
CC given in 114011-7 to the amplified sequence. These probes were
CC the respiratory tract. This spacer region is derived from Pseudomonas
CC putida LMG 2232.
SQ Query Match 1.38; Score 14; DB 21; Length 468;
SQ Best Local Similarity 100.00; Pred. No. 3,47+02;
SQ Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dn 338 atctatcatatgaaag 351
Dn |||||
Qy 1092 TATGATGATGAAAG 1105
RESULT 125
ID TS9835 standard; DWA: 494 BP.
AC TS9835:
AC 22-APR-1997 (first entry) 87
AC Beta-casein; Gnat; transgenic antithrombin III; factor XI; factor X;
KW TAT; serine protease inhibitor; thrombin; factor VII; factor IX; plasma
KW factor XII; mammary gland specific; hereditary AT3 deficiency; therapy:
KW acquired AT3 deficiency; heparin affinity; as.
KW Pseudomonas haitus.
KW Key Location/Qualifiers
PT misc_feature 1
PT /tag= a
PT /ncg= unspecified number of bases, joins do n a
PT /ncg= unspecified number of bases, joins do n a

```

```

True Nov 17 08:55:26 1998      US-08-487-977-9.rmg

PT 17-JAN-1992 (first entry)
DW Bovine epithelium (partial).
KW ET: growth regulation; inhibition; stimulation; ss.
OS Bos taurus.
PN W0911510-A.
PR 03-MAR-1991.
PR 03-APR-1991. US-003221
PR 13-APR-1990; US-054508.
PR 13-MAR-1991; US-083796.
PA (BREM.) BRISTOL-NTERS SQIBB:
PI where? 91-1251Gelman GD;
PI where? 91-1251Gelman GD;
DR P-FSDB: R15426.
CC New cysteine-rich growth modulating proteins, epithelins - useful
CC as inhibitors of neoplastic cell growth and to promote wound
CC healing and tissue plasticity.
CC Discovered and first published in English.
CC The epithelins appear to comprise several distinct members sharing
CC significant structural homology. Two members of the epithelin family,
CC EP-1 and EP-2, have been purified from neutral sources, and cDNAs
CC cloned from sheep and rat [Q14338], human [Q14339], bovine, murine
CC [Q14340] and chicken [Q14953].
CC ET-1 is a bifunctional growth regulator, capable of stimulating
CC the growth of some cell types while inhibiting the growth of others.
CC Bioactivity mutually antagonistic to ET-1 is growth inhibitory
CC factor, ET-B, also known as ET-2. It is growth stimulatory for
CC CC eliciting the growth stimulatory activity characteristic of ET-1 and,
CC in fact, antagonises this ET-1 activity.
CC See also Q14339-40, Q14952-53, R14328-9 and M15315-20.
CC Sequence 539 bp; 99 A; 171 C; 183 G; 106 T.

Query Match 1.38; Score 14; D3; Length 539;
Matches Local Similarity 100.0%; Prd. No. 3,476+02.
BD 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

DB 48 ccgcagactcggcc 61
|||||
|||

QY 891 CCTGCAGCTCCTGGC 904

PSTBLT 127
TD Q55054 standard; DNA; 546 BP.
AC Q55054;
CD 18-JUL-1994 (first entry)
PR 18-JUN-1994.
PW Sequence of orphan receptor DNA fragment PCR 9
PS PEG-10; Primer; Peptide ligand receptor; ss.
OS Synthetic.
PE Key
FH Location/Qualifiers
FT misc_feature 1..26

```











[illegible]

RESULT 136  
ID ID T93132 standard; DNA: 590 BP.  
GC T93332;  
CA 04-MAY-1998 (first entry)  
DE IMAGE clone 26611; gene fragment; human; chromosome 10; D10S51 marker;  
NM D10S51 marker; tumour suppressor gene; prostatic cancer; cancer therapy;  
NM melanoma; glioma; non-Hodgkin's lymphoma; cancer susceptibility;  
NM hyperplasia; cancer detection; cancer susceptibility;  
NM hyperplasia; low stage tumour; ss

PN WO9715686-A1.  
PD 01-MAY-1997.  
PT 22-OCT-1996; G02588.  
PR 23-OCT-1995; US-003840.  
PA (INCR ) IMPERIAL CANCER RES TECHNOLOGY.

[illegible]

Page 151

Tue Nov 17 08:55:26 1998  
US-08-887-977-9.jpg

CC proteins provides a synergistic increase in plant protection, and may  
CC also offer protection against a wider range of pathogens.

```

Query Match      1.34;      Score 15; DB 18; Length 653;
Best Local Similarity 100.0%;
Pred. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 314 caactttgtctca 328
QY 515 caactttgtctca 529
|||||

```

RESULT	138
ID	D736496 standard; cDNA: 666 BP.
AC	236496;
DT	DE 24-MAR-1997 (first entry)
DE	Male glutathione-S-transferase IIIC-encoding cDNA.
OR	Male glutathione-S-transferase IIIC-encoding cDNA; transgenic;
NW	Plant; increased herbicide resistance; hetero-oxyacetamide; aa.
OS	Zea mays.
FE	Key
PH	Location/Qualifiers
PT	1..666
CT	/cds
PR	/tag= a
PR	DE19501840-AL
PD	25-JUL-1996
PP	23-JAN-1995: 001840.
PR	23-JAN-1995: DE-001840.
PR	(Transcription Tag.)
P1	Hickelstein, Hahn R., Mann K., Reif H., Reinemer P;
P1	Thomszik JE;
DR	WPI: 96-34287/35.

CC DNA encoding glutathione-S-transferase III $\alpha$  of maize - used to  
PF herbicides resistance of plants to retro-oxylacetamide type  
PS Claim 2; Page 15-16; 11pp; English.  
P8 Claim 3 is a cDNA sequence derived from maize that encodes glutathione-  
S-transferase III $\alpha$  (GstIII $\alpha$ ). The gene imparts a natural resistance  
CC production of GSTIII detoxifying enzyme in vivo. The gene was  
CC incorporated into vectors under the control of the cauliflower mosaic  
virus promoter. The vectors may be used to transform Agrobacterium  
tumefaciens, which in turn may be used to infect almost any plant, so  
CC transformed include potato, sugar beet, sugar cane, cereals and  
CC tobacco.

Sequence 666 BP; 103 A; 263 C; 201 G; 99 T;  
SQ

Query Match 1.3%; Score 15; DA 267; Length 666;  
Best Local similarity 100.0%; Pred. No. 8,37e+01;

Query Match 1.3%; Score 15; DB 26; Length 666;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;

Page 127

Thu Nov 17 08:55:26 1998  
119-08-887-977-9. m

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      398 ACCGGTACGCGCA 412
|||||
RESULT 139
ID   G68531 standard; DMA; 672 BP.
DC   17-FEB-1995 (first entry)
DE   Wheat germ coding sequence.
KW   oxalic acid degradation; oxalate oxidase; selection marker;
OS   Triticum aestivum L.; wheat; oxalate resistance; plant cell selection; ss.
PE   Key
      Location/Qualifiers

```

WO9413790-A.  
23-JUN-1994.  
07-DDR-1993; FOI203.  
07-DEC-1992; FR-014721.  
PRA 07-DEC-1992; FR-014721.  
PA (SRAP) SOC NAT ELP AQUITAFNE.  
PI Grezes-Bessat B, Gilson R, Pignard A, Schneider M.  
DR Wt of gene encoding oxalic acid degrading enzyme - for plant  
P-PSDB: R56853.  
CC Use of gene encoding oxalic acid degrading enzyme - for plant  
CC conferring pathogen resistance.  
P8 Claim 5; Page 56; 64pp; French.  
CC The sequence coding for wheat germin, a protein induced during  
CC germination of wheat, is a preferred example of DNA coding for a  
CC sequence of interest. This DNA can be coupled to a  
CC sequence of interest, or coding for a protein conferring  
CC resistance to pathogens (esp. endochitinaases). Plant calli which  
CC have been successfully transformed by the sequences are selected  
CC by their ability to grow on oxalic acid-contg. medium. The wheat  
CC germin sequence was published in J Biol Chem. 264, 4996-4900 (1989)  
CC and the wheat germin sequence was published in J Biol Chem. 264, 4996-4900 (1989)  
CC sequences 571 BP.  
P8 137 T. 138 G.  
CC 137 T. 138 G.

```

Db      87 ccaggactctgtg 100
        |||||
Cp      886 ccaggactctgtg 873

```

cp 886 CCAGGACTTCTGTG 873



RESULTS	
ABSTRACT	T51180 standard; DNA. 688 BP.
ID	T51181
AC	T51181:
DT	15-OCT-1997 (first entry)
CC	Cryptosporidium parvum DNA sequence
CC	Cryptosporidium parvum DNA sequence
CC	Cryptosporidium parvum DNA sequence
KW	diarrhoea; PCR; polymerase chain reaction; ss.
PN	Cryptosporidium parvum.
PS	P09702291-A1.
PP	P25-JUN-1996: A00387.
PP	P30-JUN-1995: AA-003916.
PP	(OTHER) UNITY MICRODOC.
DR	WPI: 97-108909/10.
PT	New Cryptosporidium DNA, primers, probes - for detecting and identifying Cryptosporidium species in the environment and in
PT	Claim 1, Page 31, Sibip: English.
PS	The present sequence represents the purified and isolated DNA sequence
CC	of Cryptosporidium parvum. A method has been produced for the detection
CC	and identification of Cryptosporidium parvum by amplifying at least
CC	one specific primer or probe (see T7844, p. 7848-94). The primers
CC	are used to detect and/or identify Cryptosporidium microorganisms
CC	(such as Cryptosporidium parvum, which causes diarrhoea) by selecting a
CC	mixture of primers and probes (see T7844, p. 7848-94). The products of this
CC	mixture are then amplified using PCR and the products detected. The
CC	products are then analysed using RFLP and the products identified. This
CC	method is used to analyse samples for environmental contamination by
CC	Cryptosporidium oocysts and to diagnose Cryptosporidium infections in
CC	Sequence. 688 BP.
CC	211 A:
CC	108 G:
CC	157 G:
CC	212 F:

```
Query Match      1.3% Score 14: DB 32: Length 688;
Best Local Similarity 100.0%; Pred.No. 3.47e+02;
Matches 14: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 141  
 ID T79667 standard; DN: 689 BP.  
 AC T79667:  
 DT 04-FEB-1998 (first entry)  
 DE BRCA2 cancer susceptibility gene exon 23a.  
 KW BRCA2 cancer susceptibility gene; breast cancer;  
 OS Histiocytic prostatic cancer; colorectal cancer; Ocular melanoma;  
 K1 leukemia; human; de.  
 OS Homo sapiens.

Page 155

[illegible]

```

Query Match      1.3%   Score 14:  DB 22:  Length 702;
Best Local Similarity 100.0%;  Pred No. 3,47e+02;
Matches 14:  Conservative      0;  Mismatches 0;  Indels 0;  Gaps 0;

```

RESULT	143
AC	standard; DNA: 713 BP.
TG5198	AC
DT	28-NOV-1996 (first entry)
DE	Muscle Interleukin-12 p35 subunit gene.
PE	Bioactive fusion protein:
REMARKS	IL-12: p35 subunit;
	Cytokines; tumor; melanoma;
	fibrosarcoma; renal cell
KM	carcinoma; immunotherapy; therapy;
RM	retrovirus; viral ds.
	Mus musculua.
KEY	Location/Qualifiers
FR	/tag= a
CD5	/tag= b
FFI	signal_peptide 47..112
FI	
FII	
FI	
FI	
FI	
FI	mat_peptide 113..691

GR2307407-477-A.  
28-MAY-1997. 024453.  
25-NOV-1996. 024453.  
28-NOV-1996. GB-021961.  
28-NOV-1996. GB-021961.  
14-DEC-1995. GB-025555.  
(CANC.) CANCER RES CAMPAIGN TECHNOLOGY.  
(TUDU.) UNIV DUKE.  
ALMOWORTH, ALICE, MITCHELL PA, Stratton MR. Wooster RP.  
Nucleic acid molecules comprising part or all of the BRCA2 cancer  
susceptibility gene useful for diagnosis, prognosis or therapeutic  
treatment of cancer  
Claim 1, Fig 2, 1246P. English.  
The BRCA2 gene, the genomic organization of the BRCA2  
cancer susceptibility genes exon 20a. The nucleic acid molecules can be  
used to construct probes for screening cDNA or genomic libraries,  
sequencing positive clones obtained, and assembling the full length  
BRCA2 sequence. The BRCA2 nucleic acid molecules and proteins are  
especially for the clinical treatment of hereditary cancer therapy,  
breast cancer, ovarian prostate or colorectal cancer, ocular melanoma  
or Leukemia. In particular antisense oligonucleotides capable of  
hybridizing to the BRCA2 nucleic acid, pre-mRNA or mature mRNA are used  
so that the expression of the BRCA2 nucleic acid is reduced or  
abolished. The BRCA2 nucleic acid molecules are also useful for  
diagnosing susceptibility or predisposition to cancer in a patient.  
Nucleic acid molecules are used to design probes or primers for PCR to  
detect the presence of mutations in a sample of nucleic  
acid from a patient. The BRCA2 promoter region is useful for screening  
for mutations in the promoter region. The BRCA2 nucleic acid molecules  
control of the promoter. Antibodies are used to determine the presence,  
amount or location in a cell of a BRCA2 polypeptide or its mutant forms.  
The polypeptides are used to screen for binding partners, these are  
used to screen for substances which mimic the activity of BRCA2  
sequences. 689 EP. 204 A. 132 C.  
SEQUENCE 689 EP. 204 A. 132 C.

```

Query Match      138;   Score 14; DB 35; Length 689;
Best Local Similarity 100.0%; Pred. No. 3 474+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;

DBD    257 tgaagattatatttc 270
        |||||.....
        17 TGAAGATTATTTTC 30

RESULT 142
ID     T5246 standard; cDNA; 702 BP.
OT     07-NOV-1996
OT     07-NOV-1996 (first entry)

```

Page 126

W95245765-AL.  
07-PM-1996; 001787;  
08-FEB-1995; US-385335.  
(WHED) WHITEHEAD INST BIOMEDICAL RES.  
Lieschke GJ, Mulligan SC;  
1996-08-16; US-385335.  
P-PM-88924736.  
New DNA encoding fusion protein, esp. contg. IL-12 p35 and p40  
subunits - for treatment of established tumours or prevention of  
tumour establishment  
Example 1, Page 33-54; 18pp; English.  
Abstract (892464) of murine interleukin-12 (IL-12). It can  
be utilized in novel constructs in which the p35 coding sequence  
is joined to the IL-12 p40 coding sequence (see also 735200) via a  
DNA fragment (731393-97 and 73501-03) coding for a peptide linker  
(892464-45). The structure is incorporated into a retroviral  
vector (731398-1399) and used to transfect cells. CHSE-3, B16 or renal carcinoma  
IL-12 in transfected cells. CHSE-3, B16 or renal carcinoma  
cells secreting the IL-12 dimer can be used to reduce the size of  
established tumours and/or increase survival time, esp. in cases  
of melanoma, fibrosarcoma and renal cell carcinoma. 15p

```

Query Match      1.38;      Score 14;  DB 23;  Length 713;
Best Local Similarity 100.0%;  Pred. No. 3.47e+02;
Matches 14;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

bb 220 ttctgcactgctg 233
      |||||

```

RESULT	144
ID	G88356 standard; DWA, 751 BP.
DT	20-FEB-1995 (first entry)
DE	Wheat grain genomic clone qf-2.8.
EF	Oxalic acid degradation; oxalate oxidase; selection marker; germin:
NM	oxalate decarboxylase; pathogen resistance; plant cell selection; ss.
PR	Erlicium asitumvum
KE	Location/Qualifiers
FT	21..695
CD	*tag= a
	/product= germin
	/note= "clone qf-2.8"
NO	90413790-A.
PD	23-JUN-1994.
PP	07-DEC-1993: F01203.....















CC found on the surfaces of the proteins play a role in the function of  
CC Caps, and in particular, their interaction with their targets. The Caps  
CC are the primary structural elements of the capsid, and are responsible for  
CC the secondary structures of these molecules and stabilise their primary  
CC structures. The Cap family can be used as denaturants to uncoil secondary  
CC structures of nucleic acid, regardless of temperature, and, when they are  
CC added for the staining of DNA in gels promotes from the genes  
CC added. Capsid proteins are responsible for the stability of the capsid  
CC physiological temperatures, e.g. to minimise proteolysis or avoid  
CC incorrect folding.  
SQ Sequence 877 BP; 265 A; 151 C; 192 G; 269 T;  
Query Match 1.38; Score 14; DB 39; Length 877;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 139 gaggaactgacac 206  
|||||  
Cp 695 AGGTTTGCACAT 682  
RESULT 158  
ID V05148 standard; DNA; 903 BP.  
AC Q03120;  
DT 03-AUG-1990 (first entry)  
DE cDNA of Elmeria tenella encoding antigen GX3273  
OS Elmeria tenella; avian coccidiosis.  
FX Key  
FH Key Location/Qualifiers  
FT cds 1..231  
FT W09000403-A.  
FT 25-JAN-1990.  
FT 05-JUL-1989; U02918.  
FT 05-JUL-1989; U02918.  
FT (GENE) Genex Corp.  
FT W01: 90-051586/07.  
DR P-PSDB: R05218.  
CC Cloned gene or fragment encoding antigenic protein  
CC which binds with antibodies against avian coccidia. and  
CC Claim 6; Page 93; Fig 4; 134pp; English.  
CC An E. tenella cDNA library was screened with chicken immune serum. Two  
CC phase plaques were identified that produce coccidial antigens reactive  
CC to anti-coccidia serum. The coccidial DNA from these plaques was  
CC transcribed into complementary DNA. The cDNA was then inserted into  
CC a retroviral vector and transfected into NIH 3T3 cells. The resulting  
CC transduced cells were designated GX3271 (G93319) and GX3273  
CC (G93320). The DNA sequences encoding the GX3271 and GX3273 antigens do  
CC not show homology with each other, nor with the GX5401 (G93324) coding  
CC sequence. Plasmid expression vectors were constructed for the production

PT Quantity  
P6 Claim 3; Page 5; 11pp; Japanese.  
CC This sequence represents the DNA encoding a transcription controlling  
CC factor of the invention, designated #20. The transcription controlling  
CC factor in rice promotes transcription of the Rbgl gene. Use of the  
CC transcrip... is expected to be improved in quality and quantity.  
SQ Sequence 871 BP; 245 A; 242 C; 242 G; 240 T;  
Query Match 1.38; Score 15; DB 38; Length 871;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 514 gaggaactgacac 528  
|||||  
Cp 781 GAGGATCTGACAG 767  
RESULT 158  
ID V05148 standard; DNA; 877 BP.  
AC V05148-1998 (first entry)  
DE Nucleotide sequence of cold shock induced protein 8 (Csp8) of E. coli.  
KW Cold shock induced protein; Csp; Csp8; unind; secondary structure.  
KW stabilase; primary structure; denaturant; uncoil; temperature;  
KW DNA staining; gel; promoter; protein expression;  
OS Escherichia coli.  
FX Key  
FH Key Location/Qualifiers  
FT cds 528..743  
FT W05714575-A.  
FT 03-FEB-1998.  
FT 03-MAR-1998; 203806.  
FT 13-FEB-1998; U010332.  
FT 13-FEB-1998; U010332.  
FT (UNIPROT) UNIPROT: P04752.  
FT Echegaray J, Goldstein J, Inouye M, Jiang W, Jones P.  
FT Pollitt RS;  
DR W01: 98-129909/12.  
CC New cold-shock inducible proteins from Escherichia coli - are useful  
CC for stabilising or denaturing nucleic acid, or when labelled for  
CC staining nucleic acid in gels  
CC Disclosure; Fig 11; 39pp; English.  
CC Escherichia coli SH22 Cold shock protein induced protein 8 (Csp8) of  
CC Escherichia coli SH22 Cold shock protein induced protein 8 (Csp8) of  
CC response to a decrease of the temperature below ambient, or physiological  
CC growth temperature. The Csp8 are, however, stable at temperatures above  
CC the cold temperature at which they are induced. All the Csp8 have a very  
CC conserved aromatic residue at a conserved position. A large number of  
CC charged residues. It is possible that the aromatic residues that are

CC Human neuroblastoma cell line, hippocampus, frontal and temporal  
CC cortex and visual cortex cDNA banks were screened with a probe  
CC containing carp skeletal muscle Ca-channel cDNA. The cDNA clone  
CC p127.4.2.1.1 was sequenced and found to encode amino acids  
CC 127-421. The cDNA was then inserted into a retroviral vector and  
CC transduced into NIH 3T3 cells. The resulting cells were designated  
CC p127-4.2.1.1.1 (see Q9265). The sequence can be inserted into a  
CC eukaryotic expression vector for use in transforming suitable host  
CC cells. Cell lines producing human neuronal calcium channel proteins  
CC can be used for screening for compounds that modulate the function of  
CC antagonists. See also Q9265-Q9275.  
SQ Sequence 920 BP; 186 A; 263 C; 225 G; 239 T;  
Query Match 1.38; Score 14; DB 5; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 245 catctatgcatca 258  
|||||  
Cp 342 CATCTATGCACTA 355  
RESULT 161  
ID V05148 standard; DNA; 927 BP.  
AC Q9272;  
DT 08-MAY-1995 (first entry)  
DE IGF-1R signal peptide antisense sequence.  
DE Insulin-like growth factor-1 receptor; IGF-1R; signal peptide;  
DE antisense; oligonucleotides; differentiation; cancer therapy.  
OS Synthetic.  
OS Synthesis; cancer; ovary cancer; glioblastoma; ss.  
FN W09422486-A.  
FT 13-OCT-1994. U03314.  
FT 13-OCT-1994. U03314.  
FT 25-MAR-1993; U037257.  
FT 30-NOV-1993; U0158761.  
PA (UNIPROT) UNIPROT: P04752.  
PI Baserga R, Rubin R, Sell C;  
PT Inhibiting proliferation of cells using anti-sense  
PT oligo-nucleotide(s) - complementary to insulin-like growth factor  
PT 1 receptor; also causes differentiation and is used to treat  
PT cancer. See also: Page 35; 54pp; English.  
CC The antisense oligonucleotide given in Q92701 is complementary to  
CC codons 1 to 309 of the mature sequence of human IGF-1R (whose  
CC complete sequence is given in Q92702). Fragments of this antisense  
CC sequence may be used to inhibit the growth of IGF-1R-secreting  
CC tumour cells.  
SQ Sequence 927 BP; 179 A; 258 C; 277 G; 213 T;

CC of the GX3271 and GX3273 antigens in Z. coli. They were designated  
CC pGX3271 and pGX3273, and encode beta-galactosidase coccidial antigen  
CC fusion proteins. Also new are host cells transformed with the vectors,  
CC and the recombinant proteins. The vectors and the recombinant  
CC cells are used in a vaccine to immunise birds with the antigens.  
CC By labelling the peptides, they can be used as a type-specific probe.  
CC The AP may also be used in an assay to detect Ab against the coccidia.  
CC The Abs are used to identify transformed cells contg. the DNA.  
SQ Sequence 903 BP; 226 A; 203 C; 235 G; 235 T;  
Query Match 1.38; Score 15; DB 1; Length 903;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01; Indels 0; Gaps 0;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 32 actctgagcagtt 46  
|||||  
Cp 881 ACTCTGACAGTT 867  
RESULT 160  
ID Q9272 standard; DNA; 920 BP.  
AC Q9272-1995 (first entry)  
DE Human calcium channel 27580/14.  
DE Plasmid p1247-4.2.1.1: Ca-flux assay; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT misc\_difference 9; tag= a  
FT /note= 'undefined'  
FT misc\_difference 116..119  
FT /tag= b  
FT /note= 'undefined'  
FT misc\_difference 49; tag= c  
FT /note= 'undefined'  
FT misc\_difference 77  
FT /tag= d  
FT /note= 'undefined'  
FN EP-507170-A.  
PD 07-OCT-1992.  
PD 23-MAR-1992; 104570.  
PD 23-MAR-1992; 104570.  
PA (PARB) BAYER AG.  
PI Franz J, Ras P, Unterbeck A, Weingaertner B;  
DR W01: 92-333446/41.  
DR EP88D; R27652.  
CC Human calcium channel sub-type; useful in  
CC calcium flux assays to screen for neurone-specific calcium  
CC channel ligands  
CC Claim 2; Page 86-88; 101pp; German.



D	05-NOV-1996 (first entry)
DT	HepG2 hepatocarcinoma substance up-04lv gene.
KW	Human hepatocyte proliferation substance; rat; probe; SR-alpha promoter;
KM	Chinese hamster ovary cell; CHO; hepatitis; ds.
NC	Homo sapiens.
P8	Location/Qualifiers
PF	signal_peptide 1..66
PT	/tag= a
PT	mat_peptide 67..936
PT	/tag= b
PN	WO96093983-A1.
PD	28-MAR-1996.
PP	21-SEP-1995; JOL1904.
PA	1-SEP-1995; JOL1904.
PF	(TAIS ) TAISHO PHARM CO LTD.
PI	Nanada K, Hara H, Matsuki Y, Shindo S, Yoshimura R;
PR	WP1: 96-188445/19.
DR	FEDOS: 954317
DD	Abstract: Expression of hepatocyte proliferation substance - in
PT	eukaryotic cells, used as diagnostic reagent for hepatitis and for
PT	basic research on in vitro liver cell cultures
PT	Example 3; Page 28-29; 48pp; Japanese
PS	proliferation substance (HPS). The variant differs from the sequence
CC	given in T13396 at pos. 214: A to G change and pos. 314: T to C. These
CC	changes alter the amino acid residues at pos. 50 of the mature protein:
CC	Asp<->Ser, Thr<->Ile, Asp<->Glu. The gene was deposited from a
CC	Lambda-grijo human cDNA library using the gene as template.
CC	The human gene was cloned in plasmid pSCL to produce plasmid pSVLH which
CC	was subsequently recombinant to produce plasmid pCDM-dhfr in which the
CC	HPS gene is under control of the SR-alpha promoter. This plasmid is
CC	used to transfect HepG2 cells to produce the HPS protein. The HPS protein
CC	can be used to diagnose hepatitis
CC	Sequence 936 BP: 287 A; 167 C; 240 G; 242 T;
QY	Query Match 138; Score 15; DB 27; Length 936;
QY	Best Local Similarity 100.0%; Pred. No. 8.37e+01;
QY	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	612 tgggagatattggtg 626
QY	158 tggcgcatttcttg 172
Result 164	
ID	T13396 standard; cDNA to mRNA. 936 BP.
AC	T13396.
AC	T13396.
DT	05-NOV-1996 (first entry)
KW	HepG2 hepatocarcinoma substance HP-04l gene.
KM	Human hepatocyte proliferation substance; rat; probe; SR-alpha promoter;
NC	Chinese hamster ovary Cell; CHO; hepatitis; ds.

[illegible]

```

Query Match      1.3%  Score 14:  DB 12:  Length 927;
Best Local Similarity 10.0%  Pred. No. 3.47e+02;
Matches 14:  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Db 862 gqanucuccagcgc 875
QY 98 GCAGTTCCTCAGGC 111

RESULT 163
ID T139606 standard; RNA: 927 BP.
AC T739606;
DT 12-JUG-1996 (first entry)
DS IGf1 receptor codon 1-309
KW IGf1 receptor, cell growth, cell differentiation chamber;
NM Insulin growth factor 1 receptor; IGf1R; ss.
OS Synthetic.
PB PR M09614746-A1.
PP 12-MAY-1996; U14952
PP 12-MAY-1996; U14952
PP 12-NOV-1994; US-340732.
PA (UTR-) UNIV JEFFERSON THOMAS.
PI Abraham O, Hasegawa R, Resnicoff M;
PT (13-NOV-1994)
PT Induction of resistance to tumour cell growth - by insertion of
PT diffusion chamber congen, tumour cell culture supplemented with. e.g.
PT apoptosis inducing agents, into host
PT Claim 11: Page 78, 729,606; Pp1: English.
PT Mature insulin growth factor 1 receptor (IGf1R, see also T39608).
PT The antisense sequence acts as an apoptosis-inducing agent when
PT applied to tumour cells (prof. human melanoma, prostate, ovary,
PT embryonic, lung or smooth muscle) growing in a diffusion chamber.
PT CC Inducible, tumour cell growth is inhibited.
PT CC The antisense sequence is effective in inhibiting growth of
PT CC the chamber into a mammal. A decrease in the number of IGf1R,
PT CC brought about by the antisense RNA, causes massive cell death in
PT CC vivo.
PT CC
SEQ 927 BP: 179 A, 258 C, 277 G, 213 U;
Query Match      1.3%  Score 14:  DB 20:  Length 927;
Best Local Similarity 78.6%  Pred. No. 3.47e+02;
Matches 11:  Conservative 3;  Mismatches 0;  Indels 0;  Gaps 0;

Db 862 gqanucuccagcgc 875
QY 98 GCAGTTCCTCAGGC 111

RESULT 163
ID T13397 standard; cDNA to mRNA: 936 BP.

```

OS Homo sapiens. Location/Qualifiers  
FF Key signal\_peptide 1..66  
FT FT /crgs a  
FT FT /rsgs b  
FT mat\_peptide  
PB NG09609383-A1.  
PF 28-MAR-1996. 301804.  
PI R12752-1995. JP-228334.  
PR 22-SEP-1995. JP-228334.  
PT (TAIS ) TAISEN PHARM CO LTD.  
PFI Hanada K., Hara H., Matsuki Y., Shindo S., Yoshimura H.  
DR WP1: 96-18045/19.  
CC This is the nucleotide sequence encoding a human hepatocyte proliferation vector for expression of hepatocyte proliferation substance - in eukaryotic cells, used as diagnostic reagent for hepatitis and for basic research on in vitro liver cell cultures

CC This is the nucleotide sequence encoding a human hepatocyte proliferation library using a fragment of the corresp. rat gene as a probe (sequence not given in the specification). The human gene was cloned in plasmid CC produce plasmid pCDL4-shaf in which the HPS gene is under control of the CC alpha promoter. This plasmid is transformed into CHO cells to express the HPS protein. The HPS protein can be used to diagnose hepatitis.

CC Sequence 936 BP; 288 A; 186 C; 239 G; 243 T;  
Query Match 1.3%; Score 15; DB 22: Length 936;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ddb 612 tgggggnaattctgg 626  
Oy 158 TGCGGGTANTCTGG 172  
|||||||  
|||||||

RESULT 165  
ID Q03293 standard: DNA: 960 BP.  
DC NCBI: 1995-01-01  
DT 03-AUG-1990 (first entry)  
RW Recombinant DNA encoding human prolactin (HP)  
KW Human prolactin (HP); Plasmid dPT100; plasmid pOR720; plasmid pU5100.  
RN J030200.1  
PM J030200.1  
PD 03-JAN-1990.  
PF 25-DIC-1987; 331244.  
PP 25-DIC-1987; JP-331244. JP-315317.  
PT (SHIK.) SHIKASHIMA BOSEKI K.  
PFI  
DR WP1: 90-047987/07.



CC transfer to the womb of a recipient.  
 SQ Sequence 967 BP; 253 A; 233 G; 237 T;  
 Query Match 1.38; Score 14; DB 1; Length 961;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 298 ccgctgcttaccy 311  
 |||||  
 QY 909 cccgtgcttaccg 922

RESULT 167  
 ID 773867 standard; cDNA; 967 BP.  
 AC 773867-1998 (first entry)  
 DE cDNA encoding cotton fibre structural protein from clone 4-4.  
 KW promoter; fibre-specific; transcriptional factor; promoter;  
 KW altered phenotype; colour; melanin; indigo; ss.  
 KW Gossypium hirsutum cv. coker 130.  
 FT 1-967  
 FT mat\_peptide  
 FT /tag= a  
 FT /codon= seq; tga, aa: stop  
 FN W0964024.A2.  
 PN 07-JUN-1998; 109897.  
 PR 07-JUN-1998; US-480178.  
 PR 01-JUL-1998; EA-005572.  
 PA (CALJ) CALZENE INC.  
 WP1: 97-05325/05. Peres-Grau L, Stalker DM;  
 DR P-PSDB; W21900.  
 DR P-PSDB; W21900.  
 PT DNA construct cony. gene of interest controlled by cotton fibre  
 PT transcriptional factor g. used to produce altered phenotype cotton  
 PT protein. DNA constructs containing cotton fibre-specific transcriptional  
 CC factor promoters are useful to produce cotton fibre cells with altered  
 CC melanin (e.g. tyrosinase gene and OR438 encoded protein from  
 CC Streptomyces antibioticus) and indigo (mono-oxygenase genes possibly in  
 CC conjunction with a tryptophanase gene) are of interest. The promoters of  
 CC the cotton fibre cell are reliable and permit expression of a protein selectively  
 CC in the cotton fibre cell. The cotton fibre cell is a suitable system for  
 CC colour and durability as required. The construct and methods can also be  
 CC used for the introduction of other advantageous genes into a cotton  
 CC plant, e.g. a plant hormone. In particular, fibres from a plant producing  
 CC coloured and/or modified fibres may be used to produce yarns and/or fabrics that do not  
 CC require dyes.

SQ Sequence 967 BP; 358 A; 188 C; 202 G; 219 T;  
 Query Match 1.38; Score 14; DB 34; Length 967;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 206 cnaatctatctga 219  
 |||||  
 Cp 671 CAAAATCCTGAAAC 658

RESULT 168  
 ID 904289 standard; DNA; 983 BP.  
 AC 904289-1990 (first entry)  
 DE Sequence encoding ovine trophoblast protein-1 (OTP-1).  
 KW ovine trophoblast protein-1; bfp-1; fertility; ds.  
 PN EP-367063-A.  
 PD 9-MAY-1990.  
 PD 23-OCT-1989; 119642.  
 PD 23-OCT-1989; US-762870.  
 PA (DROB) DROB, Inc.  
 PI Roberts MR, Zankava K.  
 DR P-PSDB; R04540.  
 DR P-PSDB; R04540.  
 PT used for enhancing fertility or treating viral diseases in  
 PT mammals, esp. cattle.  
 PS Dielourea; P; English.  
 CC The bfp-1 produced from the gene may be used to promote fertility  
 CC in the bfp-1 produced from the gene may be used to promote fertility  
 CC provide transgenic animals with enhanced fertility or in  
 CC prophylactic and therapeutic treatment of other mammals.  
 CC Sequence 983 BP; 263 A; 260 C; 221 G; 239 T;  
 SQ Sequence 983 BP; 263 A; 260 C; 221 G; 239 T;  
 Query Match 1.38; Score 14; DB 1; Length 983;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 311 ccctgcttaccy 324  
 |||||  
 QY 909 cccgtgcttaccg 922

RESULT 169  
 ID 947200 standard; DNA; 1002 BP.  
 AC 947200-1994 (first entry)  
 DE Delta polymerase III holoenzyme delta' subunit holoB gene.  
 KW Enzyme; Chromosomal replicase; ss.

OS Escherichia coli. Location/Qualifiers  
 FH Key 1..1002  
 FT cds /tag= a  
 FN W0931515.A  
 PD 05-AUG-1993.  
 PF 22-JAN-1993; 006627.  
 PR 24-JAN-1993; US-826936.  
 PT OGDONE, J. W. CORNELL RES FOUND INC.  
 WP1: 93-255618/32.  
 DR P-PSDB; R40126.  
 DR P-PSDB; R40126.  
 PT Nucleotide sequences for DNA polymerase III holo enzyme sub-units  
 PT enzymes with 5 or 6 sub-units having same activity as 10 subunit  
 PT enzyme.  
 PS Example; Page 23-24; 115pp; English.  
 CC The sequence is that of the DNA polymerase III holoenzyme delta'  
 CC subunit holoB gene.  
 CC Sequence 1002 BP; 218 A; 267 C; 275 G; 242 T;  
 SQ Sequence 1002 BP; 218 A; 267 C; 275 G; 242 T;  
 Query Match 1.38; Score 14; DB 8; Length 1002;  
 Best Local Similarity 100.0%; Pred. No. 9.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 379 gcattctgaacac 392  
 |||||  
 Cp 320 CCAATCTGAAAC 307

RESULT 170  
 ID 798155 standard; DNA; 1002 BP.  
 AC 798155-1998 (first entry)  
 DE E. coli DNA polymerase III delta' subunit coding sequence.  
 KW Delta' subunit; E. coli; DNA polymerase III holoenzyme; probe; enzyme;  
 KW hybridization; PCR; ss.  
 FT 1-1002  
 FT Escherichia coli. Location/Qualifiers  
 FT CDS /tag= a  
 FT /gene= 'holB'  
 FT /product= 'DNA polymerase III delta' subunit  
 FT /note= 'cotton as given at the end of the  
 FT sequence'  
 FT /trans\_except= (pos: 721..723, aa: Ala)  
 FN U5568004.A.  
 PN 22-JUL-1994; 279058.  
 PR 22-JUL-1994; US-279058.  
 PR 22-JUL-1994; US-846936.  
 PA (CORR ) CORNELL RES FOUND INC.

PI O'donnell WR.  
 DR P-PSDB; W36071.  
 DR P-PSDB; W36071.  
 PT Isolated theta, chi, psi, delta and delta' protein sub-unit of E.  
 PT coli polymerase III holoenzyme used to make man-made enzyme of 5  
 PT subunits. The subunit is used in long chain PCR.  
 PS Claim 35: Column 19-20: 65pp; English.  
 CC This is the nucleotide sequence encoding the delta' subunit of the  
 CC E. coli DNA polymerase III holoenzyme. The sequence was isolated from  
 CC a lambda phage library using probes 798158-798159.  
 CC (798151-798155 respectively) are used to make man-made enzymes  
 CC comprising 5 or 6 subunits and potentially for use in long chain PCR.  
 CC Sequence 1002 BP; 218 A; 267 C; 275 G; 242 T;  
 SQ Sequence 1002 BP; 218 A; 267 C; 275 G; 242 T;  
 Query Match 1.38; Score 14; DB 36; Length 1002;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 379 gcattctgaacac 392  
 |||||  
 Cp 320 CCAATCTGAAAC 307

RESULT 171  
 ID 094308 standard; cDNA to mRNA; 1023 BP.  
 AC 094308-1996 (first entry)  
 DE cDNA encoding rat beta-1,3-galactosyl transferase; receptor;  
 KW beta-1,3-galactosyl transferase; glycoprotein; disaccharide; Gal-beta-1,3-GalNAc;  
 KW sialic acid; prevention; cancer metastasis; viral infection;  
 KW anti-inflammatory; nerve cell stimulation; P-F4H; P-F4H; ds.  
 FT 1-985  
 FT Katus norvegicus Location/Qualifiers  
 FT CDS /tag= a  
 FT /note= 'encodes residues 58-350 of rat Gal-beta-  
 FT 1,3GalNAc:alpha-2,3-sialic acid transferase'  
 FN J07236477-A.  
 PN 14-SEP-1995.  
 PR 28-FEB-1994; JP-029384.  
 PR (RIKA ) RIKAGAKU KENKYUSHO.  
 DR WP1: 95-347446/45.  
 DR P-PSDB; R75200.  
 DR P-PSDB; R75200.  
 PT useful in prevention of cancer metastasis, viral infection, etc.  
 CC Claim 8: Page 12-13; 15pp; Japanese.



---

```

Query Match      Score 13:  95.11; Length 1025;
               1.36;
Best Local Similarity 100.0%;
Matched Local    Pred. NO. 8.37e+01;
Matches         15:  Conservative      0:  Mismatches 0:  Indels 0:  Gaps 0:

Ob 269 gtcttactctccca 283
      ||||||

```

```

RESULT      173
ID          N91903 standard; DNA; 1028 BP.
AC          NCBI903:1900 (kfast entry)
DT          75-APR-1990 (date of last entry)
TX          2S albumin gene and some related elements
DE          2S albumin: Arabidopsis thaliana; pBK261: storage protein gene;
KW          Arabidopsis thaliana; protein coding; storage protein;
OS          Arabidopsis thaliana.
NC          NCBI903:1900
CC          Key Location/Qualifiers
           cds         432..926
           misc_feature 773..813 complementary
                        /tag= b
                        /note= site to which oligonucleotide is bound during
                               mutagenesis
OR          N908031887-A.
PE          PD
PM          05-MAY-1989.
PR          20-OCT-1989. E00944.
PS          1989-01-18.
PT          (PLAN) Plant Genetic Syst.
AU          Vandereschov JS, Krebers E, Botterman J, Lemmans J.
WJ          WPI: 89-150783/20.
RE          Recombinant DNA expression in plants
           PR          - using modified storage protein genes for expressing
           PR          heterologous poly(epsilon)s in the seeds
           CC          it encodes the epsilon 18-albumin storage protein including signal

```

```

DB      784  gtctctggcctcct 797
          |||
Query Match 1:36; Score 14; US 14; Length 1037;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

Db      873 GCGATGACAGCAACT 887
QY      1063 GCGATTGTCGCATC 1077

RESULTS 177 standard; DNA; 1040 BP.
AC      N90048;
DE      1-NOV-1989 (first entry)
KE      Neisseria gonorrhoea (FA19). Protein IA DNA
KW      Neisseria gonorrhoea (NG); NG strain FA19; DNA; Protein IA.
PB      Neisseria gonorrhoea (NG)
FH      Key Seria gonorr Location/Qualifiers
FT      misc_feature 1037..1040
          /tag= a
FT      -35_signal 11..36 b
          /tag= c
FT      -10_signal 50..55 b
          /tag= c
FT      rbs 69..73 d
          /tag= e
FT      misc_signal 81..134 d
          /tag= e
FT      misc_feature 207..210 f
          /tag= g
FT      misc_feature 214..249 f
          /tag= g
FT      cds 81..1025 g
          /tag= h
W09048(873)..A.
FN      O1-JUN-1989.
PP      23-NOV-1988; 004225.
PR      24-NOV-1987; DS-242758.
PA      (UTMC) University of North Carolina.
PE      Genbank; EMBL; Searching PF;
WPI: 95-178391/24
P-FSDB: P90039.
DR      Dr New Metseria gonorrhoeae Protein I gene sequences
FT      Used for detection of p. gonorrhoeae infection and
FT      used for identification of treatment of infection
PS      Nucleotide sequence encoding Protein IA of NG strain FA19.
CC      Disclosure: fig 3, 7Opp; English.
CC      CC Nucleotide sequence encoding Protein IA of NG strain FA19.
CC      See also P90039. Misc: feature (a) is SAUHI site, (c) is signal
CC      sequence, (f) is SAUHI site, and (g) is TagI site (see N90035 -
CC      N90048).
SQ      Sequence 1040 BP: 277 A; 286 C; 253 G; 224 T;
                                     138;
Query Match          100.0%; Score 15; DB 1; Length 1040;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

NM1269.
AC   27-MAR-1980      (first entry)
DE   nucleotide sequence of exon 1 from genomic clone lambda sPLA2-6 of human
DE   synovial phospholipase 2 (sPLA2) type A
DE   KW   human synovial phospholipase A2 gene; clone lambda sPLA2-6; exon 1.
DE   FX   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cast_signal     889..893
FT                   /cstage=a
FT   tata_signal     /cstage=b
FT                   /cstage=c
FT   exon            1018..1035
FT                   /base=Exon 1'
Y08901771.A.
AC   09-MAR-1989
PD   PF 23-AUG-1989; 002896.
PR   16-AUG-1989; US-231805; US-098883.
FR   16-AUG-1989; US-231805; US-098883.
FX   Johnsen LS, Salihpour JZ, Futschuh W, Vada P;
FX   WP1; 89-085194/11
DR   Mammalian synovial phospholipase A2 - used in food processing
PT   Design and screening of inflammation inhibitors, as an anticancer
PT   drug.
PS   Figure 7; 7opp; English.
CC   EMBL3-human leukocyte genomic library was screened using labelled probes
CC   (n90845 and n90897) based on sPLA2. Clone lambda sPLA2-6 is one of the
CC   clones in the library. The other four are in lambda sPLA2-DNA4 (n91259). There are five
CC   exons in lambda sPLA2-6. This is the first one.
CC   n91260.
SQ   Sequence 1080 BP; 306 A; 242 C; 303 G; 229 T;
Query Match          1.3%; Score 15; DB 1; Length 1080;
Best Local Similarity 100.0%; Prd. No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db                  41 caaacgaccacagc 55
Cp                  759 CACAGCACACCACG 745

RESULT 180
ID   DT7350 standard; DMM; 1087 BP.
NC   127350.1996      (first entry)
KW   Meloidoeyne resistant tomato acid phosphatase 1(1) (Apsl1) gene.
KW   Acid phosphatase 1; Apsl gene; Meloidoeyne; tomato; detection;
KW   determination; genotype; resistance; sensitivity; amplification;
KW   Lysosomal polyamine catabolic reaction; agricultural; hybridization; da.
OS   Lycopersicon

```



PI (SNOW) SNOW BRAND MILK PRODUCTS.  
 DR Morinaga T, Yasuda N, Higashio K, Tamaoki T;  
 WP1: 90-161480/48.  
 PR 25-SEP-1994.  
 DE cDNA for wild type human L-asparaginase gene  
 KW Leukemia; lymphoma; human; ds.  
 OS Synthetic.  
 PT Protein expression in animal cells  
 PS Claim 2: Page 21; 36pp; Japanese.  
 CC The present sequence is capable of combining with  
 CC the enhancer from alpha-fetoprotein gene of the  
 CC Large amounts of the protein can be produced. See also Q07005.  
 SQ Sequence 1091 BP; 300 A; 302 C; 233 G; 236 T;  
 Query Match 1.38; Score 14; DB 1; Length 1091;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 106 ttcgactccagta 119  
 QY 1062 CCGAGATACGACA 1075

RESULT 182  
 ID Q77819 standard; cDNA; 1093 BP.  
 AC Q77819;  
 DE 17-MAY-1995 (first entry)  
 KW Hepatic parenchymal cell growth factor; HPCF; liver diseases;  
 KW Liver cancer; cirrhosis; ss.  
 OS Homo sapiens.  
 PT Hepatic parenchymal cell growth factor peptide - is isolated  
 PT from human or animal liver cell or produced by recombinant  
 PT techniques and used for therapy of liver diseases  
 CC Q77819; Page 34; 7pp; Japanese; parenchymal cell growth factor (HPCF).  
 CC The protein or the 8-terminal peptide (R65757) may be used in the  
 CC diagnosis and treatment of liver diseases, such as liver cancer and  
 CC cirrhosis. 1093 BP; 325 A; 195 C; 276 G; 297 T;  
 SQ Sequence 1093 BP; 325 A; 195 C; 276 G; 297 T;

DT 05-JUN-1998 (first entry)  
 DE cDNA for wild type human L-asparaginase homologue.  
 KW Leukemia; lymphoma; human; ds.  
 OS Synthetic.  
 PT Protein expression in animal cells  
 PS Claim 1: Page 57; 7pp; English  
 CC The present sequence encodes a homologue of the wild type human  
 CC L-asparaginase.  
 CC A mammalian polypeptide having L-asparaginase activity, can be used  
 CC to treat malignant tumours, leukemia and lymphoma. 200 T;  
 SQ Sequence 1095 BP; 191 A; 347 C; 337 G;  
 Query Match 1.38; Score 14; DB 39; Length 1095;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 247 ttcgactccagta 260  
 QY 6 TTCGACTCCAGTCA 19

RESULT 185  
 ID V15792 standard; cDNA to mRNA; 1095 BP.  
 AC V15792;  
 DE 05-JUN-1998 (first entry)  
 KW Leukemia; lymphoma; human; ds.  
 OS Synthetic.  
 PT Protein expression in animal cells  
 PS Claim 1: Page 57; 7pp; English  
 CC The present sequence encodes a homologue of the wild type human  
 CC L-asparaginase.  
 CC A mammalian polypeptide having L-asparaginase activity, can be used  
 CC to treat malignant tumours, leukemia and lymphoma. 200 T;  
 SQ Sequence 1095 BP; 191 A; 347 C; 337 G;  
 Query Match 1.38; Score 14; DB 39; Length 1095;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 247 ttcgactccagta 260  
 QY 6 TTCGACTCCAGTCA 19

PI (SNOW) SNOW BRAND MILK PRODUCTS.  
 DR Morinaga T, Yasuda N, Higashio K, Tamaoki T;  
 WP1: 90-161480/48.  
 PR 25-SEP-1994.  
 DE cDNA for wild type human L-asparaginase gene  
 KW Leukemia; lymphoma; human; ds.  
 OS Synthetic.  
 PT Protein expression in animal cells  
 PS Claim 2: Page 21; 36pp; Japanese.  
 CC The present sequence is capable of combining with  
 CC the enhancer from alpha-fetoprotein gene of the  
 CC Large amounts of the protein can be produced. See also Q07005.  
 SQ Sequence 1091 BP; 300 A; 302 C; 233 G; 236 T;  
 Query Match 1.38; Score 14; DB 1; Length 1091;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 106 ttcgactccagta 119  
 QY 1062 CCGAGATACGACA 1075

RESULT 182  
 ID Q77819 standard; cDNA; 1093 BP.  
 AC Q77819;  
 DE 17-MAY-1995 (first entry)  
 KW Hepatic parenchymal cell growth factor; HPCF; liver diseases;  
 KW Liver cancer; cirrhosis; ss.  
 OS Homo sapiens.  
 PT Hepatic parenchymal cell growth factor peptide - is isolated  
 PT from human or animal liver cell or produced by recombinant  
 PT techniques and used for therapy of liver diseases  
 CC Q77819; Page 34; 7pp; Japanese; parenchymal cell growth factor (HPCF).  
 CC The protein or the 8-terminal peptide (R65757) may be used in the  
 CC diagnosis and treatment of liver diseases, such as liver cancer and  
 CC cirrhosis. 1093 BP; 325 A; 195 C; 276 G; 297 T;  
 SQ Sequence 1093 BP; 325 A; 195 C; 276 G; 297 T;

DT 05-JUN-1998 (first entry)  
 DE cDNA for wild type human L-asparaginase homologue.  
 KW Leukemia; lymphoma; human; ds.  
 OS Synthetic.  
 PT Protein expression in animal cells  
 PS Claim 1: Page 57; 7pp; English  
 CC The present sequence encodes a homologue of the wild type human  
 CC L-asparaginase.  
 CC A mammalian polypeptide having L-asparaginase activity, can be used  
 CC to treat malignant tumours, leukemia and lymphoma. 200 T;  
 SQ Sequence 1095 BP; 191 A; 347 C; 337 G;  
 Query Match 1.38; Score 14; DB 39; Length 1095;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 247 ttcgactccagta 260  
 QY 6 TTCGACTCCAGTCA 19

RESULT 185  
 ID V15792 standard; cDNA to mRNA; 1095 BP.  
 AC V15792;  
 DE 05-JUN-1998 (first entry)  
 KW Leukemia; lymphoma; human; ds.  
 OS Synthetic.  
 PT Protein expression in animal cells  
 PS Claim 1: Page 57; 7pp; English  
 CC The present sequence encodes a homologue of the wild type human  
 CC L-asparaginase.  
 CC A mammalian polypeptide having L-asparaginase activity, can be used  
 CC to treat malignant tumours, leukemia and lymphoma. 200 T;  
 SQ Sequence 1095 BP; 191 A; 347 C; 337 G;  
 Query Match 1.38; Score 14; DB 39; Length 1095;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 247 ttcgactccagta 260  
 QY 6 TTCGACTCCAGTCA 19







[illegible]

Tue Nov 17 08:55:26 1998

```

FF FF misc_feature /product= VP0
FF FF          /tag=c
FF FF /tag=vP3
FF FF /product= VP3
FF FF          /tag=d
FF FF /product= VP1
FF US767892B-A.
PP P-PMO RUS056/SI.
PP P-PMO RUS056/SI.
PP Repatitis A virus isolates and DNA - used to prepare vaccines for preventing hepatitis A virus infection.
PP Diclosure: P3; 3; Jppr English section of cytomolgue HAV isolate CC C57/JM-55. The isolate was obid from the stool of a cynomolgus macaque inoculated i.v. with a 10x liver homogenate obed from a macaque imported from Indonesia that spontaneously obedd HAV from its feces after being immunized against HAV by intramuscular PCR. The prod was sequenced analyzed and compared with the 2.33 kb PCR region nucleotides from human HAV isolate BH17S (Cohen, J.I. et al. (1987) Proc. Natl. Acad. Sci. USA 84, 2497-2503). The sequence, and proteins and peptides encoded by it can be used as antigenic determinants for the development of HAV vaccine. See also Q11580.
PP Sequence ID#1590      328 A:    213 C:   333 G:   378 T:
PP     Query Match       1.39%   Score 14, DB 3; Length 1152;
PP     Best Local Similarity 100.0%; Prd Mo 3_47e+02;
PP     Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;
db 875 pncanvacatcitra 888
db 348 TCSCCATACACTTGA 361

```

RESULT	193
ID	Q41458 standard; DNA: 1161 BP.
ID	Q41458
NC	30-459, 1993 (first entry)
DE	30 kD Borrelia burgdorferi protein.
OS	Borrelia burgdorferi
TI	Viralant: Lyme disease; antigen: ss.
VR	Reagent: anti-Borrelia burgdorferi
PR	Location/Qualifiers
FT	cds
FT	31...161
FT	-35_signal
FT	273..277a

```

Fr cds /*tag= a
Fr 1042..1132
Fr /*tag= b
Fr 954..960
Fr /*tag= c
Fr tata_signal
Fr W0305164.A.
Fr PD 18-MAR-1993.
Fr PF 02-SEP-1992. G01602.
Fr PP 01-SEP-1992. G01602.
Fr PA (TITLE) UNIV LEICESTER.
Fr DR Draper J.
Fr DI WPI: 93-100986/12.
Fr DP DPMAS: R11258
Fr DR transformation of marker gene and callus-specific promoter - for
Fr DR transformation of plants e.g. asparagus following wounding and/or
Fr DR callus formation
Fr PS Claim 3, Fig 5; 9ppp; English.
Fr CC generated by inverse polymerase chain reaction (PCR) using primers
Fr CC Q37065-66). The AopRI gene is a wound or pathogen induced gene. It
Fr CC has limited homology to a class of PRI-type genes in that it is
Fr CC quickly upregulated and highly active in wounded plant material both
Fr CC in Arabidopsis and in other plant species. The AopRI gene is a wound
Fr CC expression persists for several days post-wounding. The majority of
Fr CC cells in growing plants do not express the AopRI gene, therefore if
Fr CC the AopRI promoter is used to drive a marker gene in a transgenic
Fr CC plant, only a minimal amount of unwanted marker gene product will be
Fr CC produced.
Fr CC Sequence 1132 BP: 353 A; 233 C; 195 G; 351 T;
Fr 50
Fr Query Match 138; Score 14; DP 6; Length 1112;
Fr Query 14; Similarity 100; P-adj. 0; Gaps 0;
Fr Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Fr
Fr DD 730 tctaaagagcaca 743
Fr Q7 1111111111111111
Fr 712 tctaaagagcaca 725
Fr
Fr RESULT 192
Fr AC 131 standard; cDNA: 1152 BP.
Fr AC Q13697;
Fr DE 17-MAR-1992 (first entry)
Fr DE DNA encoding capsid region of cyno-HAV isolate CT-55/JN-55.
Fr DE Repatitis A virus; Cytomegic; HAV; monkey; vaccine; macaque; aa.
Fr DE 100 amino acid monomer; monomer; monomer; monomer; monomer; monomer;
Fr FF Key Location/Qualifiers
Fr FF cds 1..1152
Fr FF misc_feature 1..152
Fr FF misc_feature 1..152

```

Nov 17 08:55:26 1998

```

FF      /*tag= b
FF      296 .301
FF      /*tag= c
FF      320 .325
FF      /*tag= d
FF      zba
FF      WQ9108286-A.
FD      29-APR-1993.
FD      U09145.
FF      21-OCT-1992; US-781355.
FR      22-OCT-1991; US-781355.
FR      22-OCT-1991; US-781355.
FR      Barbour AG, Norris RMZM.
FR      WPI: 93-152478/18.
DR      P-PEDS: R36675.
DR      DR segment encoding 30 kD Borrelia burgdorferi antigenic
DR      DNA sequence from strain B312 isolated from Georgia
DR      burgdorferi-associated disease, e.g., Lyme disease
PS      Claim 1, Fig 5; 88pp; English.
CC      The 30 kD protein was identified in low-passage virulent strains of
CC      the 30 kD protein was partially sequenced and no relevant
CC      strains. The 30 kD protein was synthesized and used to identify DNA
CC      oligonucleotide probe was synthesised and used to identify DNA
CC      encoding the 30 kD protein in a low-passage virulent B. burgdorferi
CC      DNA library.
SC      Genbank accession number: AF011111.
SC      Sequence 1161 BP: 519 A; 137 C; 168 G; 337 T;

Query Match          1:3%; Score 14; DB 7; Length 1161;
Best Local Similarity 100.0%; Pred. No. 3.47e-02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db       551 CCANATACCA 564
Qy       531 CCNANNATACCA 544

RESULT 194
ID       ID# T79685
ID       ID# T79685
ID       Q# PES-1998 (first entry)
DE       BRCA2 cancer susceptibility gene exon 16 and intron boundaries.
KW       BRCA2 cancer susceptibility gene; breast cancer; ovarian cancer;
KW       prostate cancer; colorectal cancer; ocular melanoma;
KW       leukemic human; aa.
KW       Homo sapiens.
FH       Key Location/Qualifiers
FH       key 1..480
FH       Intron 1..480
FF      /number= 15
FF      /note= End of intron 15'.
FF      481..684
FF      /number= 16
FF      exon

```











Db 645 tctccagctattt 658  
 QY 103 tctccagctattt 116

RESULT 202  
 ID V00250 standard: cDNA: 1271 BP.  
 AC V00250;  
 DT 21-MAY-1998 (first entry)  
 KW human retina; ret3 partial cDNA clone G128.  
 KW cell growth; renal cell; nerve cell; renal failure; nephritis;  
 KW kidney transplant; toxic injury; hypoxic injury;  
 KW neurodegeneration; motor neuron disease; multiple sclerosis;  
 KW peripheral nerve injury; spinal cord injury; Down's syndrome;  
 KW cerebral palsy; Lyme disease; muscular dystrophy;  
 KW myasthenia gravis; tumour; therapy; ss.  
 PS Homo sapiens.  
 FT CDG Location/Qualifiers  
 FT 2..946  
 FT /\*tag= a

PP N09744356-A2.  
 PP 07-MAY-1997; U07726  
 PR 10-APR-1997; US-017427.  
 PR 08-MAY-1998; US-017427.  
 PR 07-JUN-1996; US-019300.  
 PR 18-OCT-1995; US-01859.  
 PA 1810V; BIOGEN INC.  
 PI Cate RL, Hession C, Sanicola-Nadel M;  
 DR WPI: 98-018411/02.  
 DR P-PSDB: WJ7462.  
 DR Ret3 is a retinoid acid encoding ret receptor ligands and related proteins  
 PT vectors transfected cells and antibodies used for cell  
 PT growth and improving survival of injured cells, especially renal or  
 PT nerve cells  
 PR Claim 1: Page 79-80; 11pp; English.  
 CC Ret3 is a retinoid acid encoding ret receptor ligands and related proteins  
 CC for human Ret ligand (Ret3) (see WJ7462), a key component of  
 CC the Ret signalling pathway. It was isolated from a human adult  
 CC heart cDNA library using a probe derived from a mouse ret3 partial  
 CC cDNA (see V00251). A full-length sequence (see V00251) for human  
 CC human ret3 and mouse ret3 sequences (see V00245-51) Vectors  
 CC containing ret3 DNA and prokaryotic or eukaryotic host cells  
 CC transformed or transfected with these vectors are claimed, as well  
 CC a method for production of ret3, its soluble variants and  
 CC Ret3, optionally when expressed from vectors in vivo, is used to

CC promote growth of new tissue and survival of damaged tissue,  
 CC particularly kidney or neural tissue. Typical applications are in  
 CC cell culture, neurodegeneration, motor neuron disease, multiple  
 CC sclerosis, bacterial, viral or prion infections (e.g. meningitis,  
 CC myelopathy associated with HIV or Creutzfeldt-Jakob disease),  
 CC as Down's syndrome or spinal cord injury, developmental disorders such  
 CC as peripheral neuropathy, muscular dystrophy, and the like.  
 CC peripheral nervous system (Lyme disease, muscular dystrophy and  
 CC myasthenia gravis). Fusion proteins are used to deliver toxins  
 CC etc. to Ret-expressing cells, especially tumours.  
 SQ Sequence 1271 BP; 264 A; 414 C; 316 G; 277 T;  
 Query Match  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1059 gqccagagagac 1072  
 Cp 1020 gqccagagagac 1007

RESULT 203  
 ID T94538 standard: cDNA to mRNA; 1273 BP.  
 AC T94538; 1998 (first entry)  
 DT 03-OCT-1995  
 KW Fescue; cinnamoyl CoA reductase coding sequence  
 KW lignin biosynthesis; da.  
 KW Festuca sp.  
 PS CDG Location/Qualifiers  
 FT 66..1004  
 FT /\*tag= a  
 FT /product= cinnamoyl CoA reductase

PP N0712893-A1.  
 PP 03-OCT-1995; P01544.  
 PR 03-OCT-1995; FR-011623.  
 PA (CHRS) CEST NAT RECH SCI.  
 PA (CHRS) INST NAT RECH AGRONUMIQUE.  
 WP1: 97-226225/20. AM, Briat JF, Gamas P, Grima-Pettenati J;  
 DR P-PSDB: W36044.  
 DR Transformation of plants to reduce or increase lignin content -  
 PT using DNA encoding alfalfa or maize cinnamoyl CoA reductase or the  
 PT corresponding anti-sense mRNA, e.g. to improve digestibility of  
 PT fodder crops  
 PS Disclosure: Page 51-54; 86pp; French.  
 CC This is the nucleotide sequence encoding the fescue cinnamoyl CoA  
 CC reductase (CCR). Note: this sequence is specifically excluded from the  
 CC claimed sequences in the specification, which are the alfalfa and maize

CC CCR genes (T94534-5). The alfalfa and maize CCR genes can be used to  
 CC transform plant cells to create transgenic plants in which lignin  
 CC biosynthesis is regulated so that lignin content is higher or lower  
 SQ Sequence 1273 BP; 286 A; 363 C; 369 G; 245 T;  
 Query Match  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 640 tgaacctgtgtc 653  
 QY 905 tgaacctgtgtc 918

RESULT 204  
 ID T62570 standard: DNA; 1276 BP.  
 AC T62570;  
 DT 27-OCT-1997 (first entry)  
 KW Heat-resistant barley beta-amylase gene promoter.  
 KW Barley; heat resistant; promoter; genetic engineering; transgenic;  
 KW plant; improved maturation; da.  
 PS N09702353-1998.  
 PD 23-JAN-1997.  
 PF 05-JUL-1996; J01866.  
 PR 05-JUL-1995; JP-191028.  
 PI 12-R; Kihara M, Okada Y, Yoshigi B;  
 DR WPI: 97-108966/10.  
 PR Expression promoter for genes inserted into plant seeds - e.g.  
 PT heat-resistant beta-amylase gene inserted into barley seeds  
 PT heat-resistant beta-amylase gene inserted into rice seeds  
 CC (EC 3.2.1.2) gene. The promoter is used for the production of  
 CC transgenic barley plants and seeds. Such plants and seeds have  
 CC improved maturation properties.  
 SQ Sequence 1276 BP; 393 A; 237 C; 234 G; 412 T;  
 Query Match  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 327 actctctctcaca 340  
 Cp 994 actctctctcaca 981

RESULT 205  
 ID T94510 standard: DNA; 1296 BP.  
 AC T94510;

DT 12-FEB-1998 (first entry)  
 DE Toxoplasma gondii p28 genomic sequence.  
 KW Toxoplasma gondii; p28; tachyzoite; probe; hybridisation; detection;  
 OS Toxoplasma gondii; tachyzoite; immunosay; antibody; toxoplasmosis; ss.  
 FH Key Location/Qualifiers  
 FH 171..1174  
 FT CDG  
 FT /\*tag= a  
 FT /product= p28 protein

PT Exon  
 FT 521..765  
 FT Intron  
 FT /\*tag= b  
 FT /number= 1  
 FT Exon  
 FT 768..1171  
 FT /\*tag= c  
 FT /number= 1  
 FT Exon  
 FT 768..1171  
 FT /\*tag= d  
 FT /number= 2

US565542-A  
 DT 09-SEP-1997;  
 PP 03-FEB-1992; 264951.  
 PR 03-NOV-1999; US-431578.  
 PR 03-FEB-1992; US-264951.  
 PI De Araujo PC, Prince JB, Ramington JS, Sharma SD;  
 DR WPI: 97-456759/42.  
 DR P-PSDB: W36012.  
 DR Nucleic acid encoding Toxoplasma gondii p28 protein - used in  
 PT diagnosis of Toxoplasma gondii infection and diagnostic reagents  
 PT expressing proteins useful in diagnosis and diagnostic reagents  
 PS Claim 1: Column 19-20; 14pp; English.  
 CC This nucleotide sequence corresponds to the genomic sequence encoding  
 CC the Toxoplasma gondii p28 protein. The sequence was obtained from a  
 CC cDNA library generated from nucleic acid extracted from T. gondii tachyzoites.  
 CC This sequence represents a full-length cDNA of 600 bp.  
 CC library. Fragments of the sequence e.g. T94511.5 can be used as probes  
 CC in hybridisation tests for direct detection of T. gondii infection i.e.  
 CC toxoplasmosis. The encoded protein is useful in vaccines and diagnostic  
 CC tests. The sequence is deposited with the EMBL/GenBank/CCDS databases.  
 CC Antisense peptide from the p28 P290 C; 370 G; 301 T;  
 SQ Sequence 1296 BP; 335 A;

Query Match  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 921 aaagcaagatg 934  
 Cp 193 aaagcaagatg 180







```

FT      /tag="c
FT      660..902 spliced out in KU-2"
FT      /tag="d
FT      /note="deleted in sld mutant"
FT      /tag="c
FT      /note="primer.#1
FT      /note="sense"
FT      901..929 f
FT      /label="primer.#2
FT      /note="antisense"
FT      963..978 g
FT      /label="primer.#3
FT      /note="antisense"
FT      1317..1333
FT      /label="primer.#4
FT      /note="antisense"
FT      664
FT      /tag="i
FT      /note="see comments"
FT
FT WO9203459-A.
FT 05-MAR-1992.   US6130
FT 27-AUG-1991.   US-594130
FT 05-OCT-1990.   US-594106.
FT (SLOK ) SLOAN-REITERING INST.
FT
FT Bemmer, P.; Rocka, K.; Buck, J.; Moore MA:
FT New purified mammalian c-kit receptor ligands stimulating mast
FT cells - and erythroid progenitor(s), treating leukemia, allergy,
FT leukopenia, anemia, AIDS, nerve damage, and enhancing bone marrow
FT recovery. Page 69; 15pp; English
FT The sequence was obtd. from a clone isolated from a cDNA library
FT prepd. from total RNA from Balb/c 3T3 fibroblasts. It encodes the
FT c-kit receptor ligand, KU-1. The sequence was found to be "in
FT frame" with respect to the sequence of the c-kit gene. The
FT base change at position 664 which leads to the substitution of Ser
FT 206 with Ala* (sic). (The sequence given in the fig. has the codon
FT -cct- at 664-666, encoding Pro). PCR was used to investigate the
FT relationship between the c-kit receptor ligand and the c-kit receptor.
FT Genetic analysis of PCR products was performed. The sequence of the
FT two forms of KU transcript of 870 and 750 bp respectively. The
FT smaller form, designated K2, was shown to have a deletion near an
FT exon boundary, the missing exon preceding the transmembrane domain.
FT The larger transcript, shown here, synthesized as a transmembrane

```

Page 370

Matches	14:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps
Dbb	1250	tgtctgaacaccac	1263					
Cpb	316	TGCTGAACACCAC	303					
 RESULT 213								
ID	Q25180	standard; DMA; 1347 BP.						
DE	20-NOV-1992	(first entry)						
DC	Enterococcus faecium vanC	antibiotic resistance gene.						
DD	Glycopeptide antibiotic: vancomycin; teicoplanin; resistant;							
DE	D-Ala-D-Ala ligase; peptidoglycan precursor; transposon;							
DF	Enterococcus faecium BM4147.							
FF	Key	Location/Qualifiers						
FF	cda	215..3						
FF	rbs	/product= vanc						
FF		/start= b						
FF		/stop= b						
FF	W09207942-A							
FF	14-MAY-1991	FQ0855.						
FF	29-OCT-1991	FQ0855.						
FF	31-OCT-1990	FR-013579.						
FF	18-SEP-1990	WNCST PASTEUR						
FF	18-SEP-1990	WNCST PASTEUR						
FF	WPI: 92-18367/72.	Dr. P. Dutka-malen S, Molins C;						
FF	P-PSDB: R24303.							
DR	Polypeptides involved in expression of glycopeptide antibiotic							
FF	resistance genes in Enterococcus faecium strains							
FF	enterococcal strains e.g. Enterococcus faecium and E.gallinarum							
FF	Claim 1: Fig 9, 161pp:	French.						
CC	This is the sequence of the SacI-PstI fragment of plasmid pAT216							
CC	belonging to the DNA region encoding the vanC protein encoded by the vanC							
CC	gene located on the DNA region of sufficient homology to vanA to allow							
CC	The vanC gene has regions of sufficient homology to vanA to allow							
CC	the same probes to be used to detect both. Preferred compositions							
CC	of the invention comprise the vancomycin resistance genes vanH,							
CC	vanX, vanY, vanZ, vanD, vanE, vanF, vanG, vanJ, vanL, vanM,							
CC	resistance on cells. See Q25178-Q25183							
CC	Sequence 1347 BP:	300 C:	279 G:	371 T:				
SQ								
 Query Match								
Blast Local Similarity	100.0%	Prod No.	347e+02					
- Matches	14:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps
Dbb	1017	ccagcagaataggag	1030					
Cpb	145	CACAGATAGAG	132					



[illegible]

Antigen: Immunogen: vaccine; tuberculosis; non specific adjuvant;	
Accession: 1362 BP	
OS: Mycobacterium tuberculosis; ss.	
Key: Mycobacterium tuberculosis; ss.	
Location/Qualifiers	
CD5	88..1122
FF	/tag=
PD	/product= Antigen_TB8A17
PD	W09709428-A2.
PD	13-MAR-1997.
PP	30-AUG-1996; 014674.
PP	01-SEP-1995; US531674.
PP	01-SEP-1995; US531674.
PP	22-SEP-1995; US533634.
PP	22-MAR-1996; US620874.
PP	22-MAR-1996; US620883.
PP	(COORD.) CORTEX CORP.
PP	Campo-neto-A, Dillon DC, Houghton R, Reed SG, Skealy YAW;
PP	Trudeau DR, Vedvick TH;
PP	1997; 322603/17.
PP	new immunogenic polypeptide(s) from Mycobacterium tuberculosis - also
PP	useful in vaccines for prevention or treatment of tuberculosis, are
PP	for diagnosis 53-53, 18990; English.
PP	A new immunogenic polypeptide has been developed comprising an
PP	immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
CC	its variant differing only in conservative substitutions and/or
CC	modifications). The present sequence encodes a specifically claimed
CC	protein containing one or more of the proteins or one of the proteins
CC	plus ESAT-6, are useful in vaccines, preferably when formulated with
CC	non-specific adjuvant, to induce an immune response against
CC	Mycobacterium tuberculosis; treatment of 512 C; 173 T;
CC	Sequence 1362 BP; 198 A.
CC	Query Match
CC	Local Similarity 1.38; Score 14; DE 34;
CC	Match 14; Conservative 0; Mismatches 0; Indels 0; Gaps
DB	1196 cagtggaagctgc 1209
OT	
OT	597 CAGGTGGAGCTGC 610
RESULT 218	
CD	13-MAR-1997 standard; DNA; 1362 BP.
AC	T31406
DT	12-JAN-1998 (first entry)
DE	Mycobacterium tuberculosis antigen TB8A17 encoding DNA.
DE	Antigen: Immunogen: vaccine; tuberculosis; non specific adjuvant;



OS Mycobacterium tuberculosis  
 PH CDS  
 FT 88 1122  
 FT /tag= a  
 FT /product= Antigen\_Tbaal7  
 FT  
 FT W0970428-22  
 PD 134461994  
 PD 30-AUG-1996: U14675  
 PR 12-JUL-1996: US-680573  
 PR 01-SEP-1995: US-521453  
 PR 22-MAR-1995: US-521280  
 PR 05-JUN-1996: US-658800  
 PA (CORI-) CORIXA CORP.  
 PI Campos-neto A., Dillon DC, Houghton R., Read SG, Skeiky YAW.  
 DR WPI: 97-192904/21  
 DR P-PSDB: W32357  
 DR  
 PT New immunogenic polypeptide(s) from soluble M. tuberculosis antigens  
 PT Useful for diagnosis of M. tuberculosis infection  
 CC A new immunogenic polypeptide has been developed comprising an  
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or  
 CC its variant differing only in conservative substitutions and/or  
 CC modifications). The present sequence encodes a specifically claimed  
 CC immunogenic polypeptide which is useful for diagnosis of M. tuberculosis  
 CC used to diagnose M. tuberculosis infection by forming complexes with  
 CC specific antibodies in the sample. Fragments of DNA encoding the  
 CC immunogenic polypeptide can be used as diagnostic primers or probes  
 CC or as a variant mutant to the antigen, especially monoclonal antibodies  
 CC or as a variant polypeptide substituted with a tag and 173 T;  
 SQ Sequence 1362 BP; 158 A; 345 C; 355 G; 399 T;  
 Query Match 1.38; Score 14; DB 34; Length 1362;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02; Indels 0; Gaps 0;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1186 caggtggaagctgc 1209  
 QY 597 CAGAGGAGGAGCTG 610  
 RESULT 219  
 ID W09719 standard: DNA: 1377 BP.  
 AD W09719 (first entry)  
 DT 19-MAY-1998 (first entry)  
 DE CP4-EPSPS gene.  
 KW 1-aminocyclopropane-1-carboxylic acid deaminase; ACC deaminase;  
 KW fruit ripening, ethylene production; glyphosate resistance;  
 KW 5'-enolpyruvylshikimate-3-phosphate synthase; EPSPS; CP4; ss.  
 OS Synthetic.

PR 31-AUG-1990: US-576537.  
 PR 28-AUG-1991: US-748611.  
 PA (MONS ) MONSANTO CO.  
 PI Barry GF, Kishore GM, Padgett SR, Stallings WC;  
 DR WPI: 97-297418/21  
 DR New isolated 5-enol-pyruvyl:shikimate-3-phosphate synthase gene -  
 PT glyphosate herbicidal plants to produce plants which are tolerant to  
 PT glyphosate herbicidal plants to produce plants which are tolerant to  
 PS Claim 3: Column 75-76; 154pp: English.  
 CC 77314 encodes a class II 5-enolpyruvylshikimate-3-phosphate synthase  
 CC (EPSPS) enzyme, derived from Agrobacterium sp. The sequence was altered  
 CC to produce a plant with a modified EPSPS gene. The modified EPSPS gene  
 CC confers tolerance to glyphosate herbicides. The modified EPSPS gene  
 CC is located in a plant cell region that could function as a polyadenylation  
 CC sites or potential RNA destabilization region. Class II EPSPS enzymes  
 CC are tolerant to glyphosate herbicides. EPSPS and sequences encoding it  
 CC are found in a wide variety of plants including corn, soybean, cotton, alfalfa,  
 CC plants such as corn, wheat, rice, barley, soybean, cotton, sugarbeet,  
 CC oilseed rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa,  
 CC poplar, pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf  
 CC grasses.  
 SQ Sequence 1377 BP; 278 A; 345 C; 355 G; 399 T;  
 Query Match 1.38; Score 14; DB 32; Length 1377;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02; Indels 0; Gaps 0;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 ccacaggtcctcc 108  
 CP 978 CCACAGGAGCTGCTCA 965  
 RESULT 221  
 ID Q23531 standard: DNA: 1377 BP.  
 AC Q23531: 992 (first entry)  
 DE Synthetic Class II EPSPS gene.  
 KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;  
 KW 5'-enolpyruvylshikimate-3-phosphate synthase; glyphosate resistant  
 OS Synthetic.  
 PD 19-MAR-1992  
 PD 28-AUG-1991: U06148.  
 PR 31-AUG-1990: US-576537.  
 PA (MONS ) MONSANTO CO.  
 DR WPI: 97-114367/21  
 DR DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate  
 PT synthase - for producing plants and bacteria tolerant to  
 PT glyphosate herbicides

PH Key  
 FT CDS  
 FT 3...1370  
 FT /tag= a  
 FT /product= CT4-EPSPS  
 FT  
 FT US5702933-A.  
 PD 30-DEC-1997: 553843  
 PD 07-NOV-1995: US-528457  
 PR 26-DEC-1990: US-632440  
 PR 06-NOV-1995: US-653943  
 PA (MONS ) MONSANTO CO.  
 PI (MONS ) MONSANTO CO.  
 DR WPI: 98-074439/0  
 DR P-PSDB: W39426  
 DR  
 PT Production of plants with delayed ripening - using DNA encoding  
 PT 1-aminocyclopropane-1-carboxylic acid deaminase  
 PT 1-aminocyclopropane-1-carboxylic acid deaminase  
 CC This sequence encodes the synthetic CP4 5-enolpyruvyl-3-shikimate  
 CC phosphate synthase (EPSPS) gene which is capable of conferring  
 CC resistance to glyphosate and is used in a novel method for producing  
 CC fruit-bearing plants with delayed ripening. The method involves the  
 CC transformation of a plant cell with a DNA construct encoding the EPSPS  
 CC gene in a plant at a level sufficient to reduce ethylene production in  
 CC the fruit.  
 SQ Sequence 1377 BP; 278 A; 344 C; 355 G; 400 T;  
 Query Match 1.38; Score 14; DB 39; Length 1377;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02; Indels 0; Gaps 0;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 ccacaggtcctcc 108  
 CP 978 CCACAGGAGCTGCTCA 965  
 RESULT 220  
 ID 77314 standard: DNA: 1377 BP.  
 AC 77314:  
 DT 02-OCT-1997 (first entry)  
 DE EPSPS DNA having sequence altered for increased expression in plants.  
 KW 1-aminocyclopropane-1-carboxylic acid deaminase; EPSPS; shikimate;  
 KW glyphosate resistant; transgenic plant; herbicide; shikimate;  
 KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;  
 KW tobacco; da.  
 OS Synthetic.  
 Location/Qualifiers  
 FT CDS  
 FT 3...1370  
 FT /tag= a  
 FT US5633435-A.  
 PD 31-AUG-1990: 576537.  
 PR 31-AUG-1990: 576537.

PS Disclosure: Fig 8: 14pp: English.  
 CC The sequence is that of a synthetic Class II 5'-enolpyruvylshikimate-3  
 CC phosphate synthase enzyme (EPSPS) gene. It can be used to create  
 CC glyphosate resistant plants or seeds by transforming plant cells.  
 CC The transformed cells are then grown into plants which are tolerant to  
 CC weeds. The crops selected for are e.g. corn, wheat, rice, oilseed  
 CC rape, tobacco and alfalfa. This provides a cost effective,  
 CC environmentally compatible weed control device.  
 SQ Sequence 1377 BP; 278 A; 345 C; 355 G; 399 T;  
 Query Match 1.38; Score 14; DB 3; Length 1377;  
 Best Local Similarity 100.0%; Pred. No. 3.47e-02; Indels 0; Gaps 0;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 ccacaggtcctcc 108  
 CP 978 CCACAGGAGCTGCTCA 965  
 RESULT 222  
 ID 793791 standard: DNA: 1377 BP.  
 DC 173791: 1998 (first entry)  
 DE Synthetic CP4 Class II EPSP synthase (EPSPS) gene.  
 KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;  
 KW shikimate acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;  
 KW 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;  
 OS Synthetic.  
 PD 06-MAY-1997.  
 PD 06-MAY-1997.  
 PR 07-JUN-1995: 476008  
 PR 31-AUG-1990: US-576537.  
 PR 28-AUG-1991: US-749611.  
 PA (MONS ) MONSANTO CO.  
 DR Barry GF, Kishore GM, Padgett SR, Stallings WC;  
 DR WPI: 97-297418/21  
 PT Production of glyphosate-herbicide tolerant plants - using DNA  
 PT encoding class II 5'-enolpyruvyl:shikimate-3-phosphate synthase  
 PT enzyme(s)  
 CC Disclosure: Fig 8: 15pp: English.  
 CC This sequence encodes the synthetic glyphosate-tolerant  
 CC 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene. It is based on  
 CC the novel EPSPS gene from Agrobacterium sp. strain CP4, which produces an  
 CC EPSPS enzyme that has little homology with known Class I EPSPS enzymes,  
 CC and belongs to a new class, Class II. The EPSPS enzymes are part of the  
 CC shikimate acid pathway. The EPSPS enzymes are involved in the biosynthesis of aromatic  
 CC compounds. EPSPS converts phosphoenolpyruvic acid (PEP) and 3-  
 CC 3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is



















CC The vectors are used in a method for evaluating the safety of a chemical compound which is a chemical compound with recombinantly produced human prothymocyte leukemia cells (RPMI-8226) and a variant (Q8716-12), or 344 (Q8717) or their auxiliary species and variants (Q8718-12), and yeast NADPH-P450 reductase, either as a fused protein or in cell extracts, and analyzing the resulting metabolite to assess the carcinogenic or mutagenic form through metabolism in the liver. Whether the chemical compound or its metabolite is useful for determining a carcinogenic or mutagenic form through metabolism in the liver.

Query Match 1.38; Score 15; DB 15; Length 1473;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1123 agtctctctctgac 1137  
QY 937 AGTCTCTCTCTGAC 951

## RESULT 239

ID T33260 standard; cDNA; 1481 BP.  
AC T33260; 1996 (first entry)  
DE RAR-alpha DNA sequence.  
DE Hammerhead ribozyme: acute promyelocytic leukemia; APL;  
KW external guide sequence: BGS; antisense: as.  
OS Homo sapiens.  
PN W09618713-A2.  
PP 14-DEC-1994; US-354956.  
PA (INO-1) INNOVIR LAB INC.  
DI George S7100; Goldberg AB. Pace U;  
PT RNA construct(s) including ribozyme(s) and antisense oligonucleotide(s) for the inactivation of RNA associated with, e.g. promyelocytic leukemia or follicular lymphoma  
Example 1: Page 55-56; 81pp; English  
CC Retinoic acid receptor (RAR-alpha) RNA (T33247) is a substrate for ribozymes, external guide sequences and antisense oligonucleotides (see also T33245 and T33246) used for the treatment of acute promyelocytic leukemia (APL). The RAR-alpha gene is located on chromosome 15 and 17, resulting in a fusion of the retinoic acid receptor gene (see also T33260) and a gene for a putative transcription factor, (T33255-57) that specifically cleave the PML-RAR-alpha fusion mRNA but not wild-type RAR-alpha mRNA (see also T33247).  
SQ Sequence 1481 BP; 309 A; 487 C; 425 G; 257 T; 3 U;

Query Match 1.38; Score 15; DB 22; Length 1481;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 245 acaagctctctgacct 239  
QY 995 ACAAGCTCTCTGACCT 1009

## RESULT 240

ID T33247 standard; RNA; 1481 BP.  
AC T33247; 1996 (first entry)  
DE RAR-alpha RNA sequence.  
DE Hammerhead ribozyme: acute promyelocytic leukemia; APL;  
KW external guide sequence: BGS; antisense: as.  
OS Homo sapiens.  
PN W09618713-A2.  
PP 14-DEC-1994; US-354956.  
PA (INO-1) INNOVIR LAB INC.  
DI George S7100; Goldberg AB. Pace U;  
PT RNA construct(s) including ribozyme(s) and antisense oligonucleotide(s) for the inactivation of RNA associated with, e.g. promyelocytic leukemia or follicular lymphoma  
Example 1: Page 55-56; 81pp; English  
CC Retinoic acid receptor (RAR-alpha) RNA (T33247) is a substrate for ribozymes, external guide sequences and antisense oligonucleotides (see also T33245 and T33246) used for the treatment of acute promyelocytic leukemia (APL). The RAR-alpha gene is located on chromosome 15 and 17, resulting in a fusion of the retinoic acid receptor gene (see also T33260) and a gene for a putative transcription factor, (T33255-57) that specifically cleave the PML-RAR-alpha fusion mRNA but not wild-type RAR-alpha mRNA (see also T33247).  
SQ Sequence 1481 BP; 308 A; 425 G; 260 U;

Query Match 1.38; Score 15; DB 22; Length 1481;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 245 acaagctctctgacct 239  
QY 995 ACAAGCTCTCTGACCT 1009

## RESULT 241

ID T44067 standard; cDNA; 1507 BP.

AC T44067:  
DE T33260 (first entry)  
DE RAR-alpha gene up-regulated in regenerating liver.  
KW hepatoma; transgenic animal; diagnosis; gene therapy; as.  
OS Homo sapiens.  
PN W09618713-A2.  
PP 14-DEC-1994; US-354956.  
PA (INO-1) INNOVIR LAB INC.  
DI George S7100; Goldberg AB. Pace U;  
PT RNA construct(s) including ribozyme(s) and antisense oligonucleotide(s) for the inactivation of RNA associated with, e.g. promyelocytic leukemia or follicular lymphoma  
Example 1: Page 55-56; 81pp; English  
CC Retinoic acid receptor (RAR-alpha) RNA (T33247) is a substrate for ribozymes, external guide sequences and antisense oligonucleotides (see also T33245 and T33246) used for the treatment of acute promyelocytic leukemia (APL). The RAR-alpha gene is located on chromosome 15 and 17, resulting in a fusion of the retinoic acid receptor gene (see also T33260) and a gene for a putative transcription factor, (T33255-57) that specifically cleave the PML-RAR-alpha fusion mRNA but not wild-type RAR-alpha mRNA (see also T33247).  
SQ Sequence 1481 BP; 309 A; 487 C; 425 G; 257 T; 3 U;

Query Match 1.38; Score 15; DB 25; Length 1507;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 68 ctctctctctgacct 82  
QY 75 CTCTCTCTCTGACCT 89

## RESULT 242

ID T33260 standard; cDNA; 1512 BP.  
AC T33260; 1996 (first entry)  
DE RAR-alpha DNA sequence.  
DE Hammerhead ribozyme: acute promyelocytic leukemia; APL;  
KW external guide sequence: BGS; antisense: as.  
OS Homo sapiens.  
PN W09618713-A2.  
PP 14-DEC-1994; US-354956.  
PA (INO-1) INNOVIR LAB INC.  
DI George S7100; Goldberg AB. Pace U;  
PT RNA construct(s) including ribozyme(s) and antisense oligonucleotide(s) for the inactivation of RNA associated with, e.g. promyelocytic leukemia or follicular lymphoma  
Example 1: Page 55-56; 81pp; English  
CC Retinoic acid receptor (RAR-alpha) RNA (T33247) is a substrate for ribozymes, external guide sequences and antisense oligonucleotides (see also T33245 and T33246) used for the treatment of acute promyelocytic leukemia (APL). The RAR-alpha gene is located on chromosome 15 and 17, resulting in a fusion of the retinoic acid receptor gene (see also T33260) and a gene for a putative transcription factor, (T33255-57) that specifically cleave the PML-RAR-alpha fusion mRNA but not wild-type RAR-alpha mRNA (see also T33247).  
SQ Sequence 1507 BP; 309 A; 487 C; 425 G; 257 T; 3 U;

Query Match 1.38; Score 15; DB 25; Length 1512;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 75 ctctctctctgacct 89

DE T33260 (first entry)  
DE T33260 (first entry)  
DE RAR-alpha gene up-regulated in regenerating liver.  
KW hepatoma; transgenic animal; diagnosis; gene therapy; as.  
OS Homo sapiens.  
PN W09618713-A2.  
PP 14-DEC-1994; US-354956.  
PA (INO-1) INNOVIR LAB INC.  
DI George S7100; Goldberg AB. Pace U;  
PT RNA construct(s) including ribozyme(s) and antisense oligonucleotide(s) for the inactivation of RNA associated with, e.g. promyelocytic leukemia or follicular lymphoma  
Example 1: Page 55-56; 81pp; English  
CC Retinoic acid receptor (RAR-alpha) RNA (T33247) is a substrate for ribozymes, external guide sequences and antisense oligonucleotides (see also T33245 and T33246) used for the treatment of acute promyelocytic leukemia (APL). The RAR-alpha gene is located on chromosome 15 and 17, resulting in a fusion of the retinoic acid receptor gene (see also T33260) and a gene for a putative transcription factor, (T33255-57) that specifically cleave the PML-RAR-alpha fusion mRNA but not wild-type RAR-alpha mRNA (see also T33247).  
SQ Sequence 1481 BP; 308 A; 425 G; 260 U;

Query Match 1.38; Score 15; DB 22; Length 1481;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 245 acaagctctctgacct 239  
QY 995 ACAAGCTCTCTGACCT 1009

## RESULT 241

ID T44067 standard; cDNA; 1507 BP.

Query Match 1.38; Score 15; DB 34; Length 1512;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 75 ctctctctctgacct 89











CC from a phase lambda phage library. DNA of sequence  
CC G5669 isolated from a human cell line. The  
CC cDNA as probe, and included the transcriptional regulatory element  
CC of tek comprising the initiation codon and untranslated sequences.  
CC The element may be used in gene therapy to introduce foreign genes  
CC into endothelial cells to correct the genetic defect.  
SQ Sequence 1599 BP; 461 A; 327 C; 391 G; 411 T;

Query Match 1.38; Score 14; DB 10; Length 1590;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 216 agcgtattttacc 229  
|||||  
QY 108 AGCGTATTTCAC 121

## RESULT 253

ID 777094 standard: cDNA to mRNA; 1593 BP.  
AC 777094-1998 (first entry)  
DE Human N-acetylglucosaminyl transferase III gene.  
KW N-acetylglucosaminyl transferase III; GAT-III; rat; human;  
KW virus; replication; inhibitor; hepatitis B; hepatitis C; HIV;  
KW viral disease; human immunodeficiency virus; da.  
PB Key septins. Location/Qualifiers  
FT CDS 1..1593  
FT /tag= a  
FT /product= N-acetylglucosaminyl transferase III  
FT /note= "No stop codon given."  
PN M09718836-1.  
PD 29-MAY-1997.  
PF 17-JUL-1996; J01886.  
PR 17-NOV-1995; JP-322474.  
PI Chiba Y, Miyoshi E, Taniguchi N;  
PI Thara Y, Miyoshi E, Taniguchi N;  
DR P-PSDB: 824015.  
DE The virus replication inhibitor containing N-acetylglucosaminyl  
DE transferase gene is useful in treatment of hepatitis B, C  
DE and HIV viral infections.  
PS Claim 6: Page 22-26; 12pp; Japanese.  
CC The present sequence represents human N-acetylglucosaminyl transferase  
CC III (GAT-III) which is a replication inhibitor of hepatitis virus  
CC (GAT-III) and is useful in the treatment of hepatitis B, C and HIV.  
CC The virus replication inhibitor is used against hepatitis B, C and HIV.  
CC It is used for the treatment of viral diseases.  
SQ Sequence 1593 BP; 269 A; 571 C; 488 G; 265 T;  
Query Match 1.38; Score 15; DB 35; Length 1593;

Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 395 agagaagctcagg 409  
|||||  
CP 1015 AGGAGAGCTTCAGG 1001

## RESULT 254

ID 729199 standard: cDNA to mRNA; 1593 BP.  
AC 729199-1996 (first entry)  
DE Rat N-acetylglucosaminyl transferase-III gene, inhibitor of metastasis.  
KW N-acetylglucosaminyl transferase-III; metastasis inhibitor; cancer;  
KW neoplasia; invasive; da.  
PB Rat sp. Location/Qualifiers  
FT CDS 1..1593  
FT /tag= a  
FT /note= "No stop codon."  
FT 708100130-A.  
PD 30-APR-1996.  
PF 12-OCT-1994; 271802.  
PR 12-OCT-1994; JP-271802.  
PI (TAKI) TAKARA SEIKO CO LTD.  
DR P-PSDB: 897615/27.  
DE Inhibitor of cancer metastasis - contains N-acetylglucosaminyl  
DE transferase-III isolated from rat kidneys  
DE Claim 3; Page 8-10; 12pp; Japanese.  
CC Claim 3 and 13: The cDNA clones of the rat N-acetylglucosaminyl  
CC transferase-III (GAT-III) gene, which is a replication inhibitor of  
CC cells and cells surrounding cancer cells for the inhibition of  
CC metastasis, preventing the spread of cancer cells to tissues other  
CC than the primary site of origin. The clones were isolated from a rat  
CC kidney homogenate.  
SQ Sequence 1593 BP; 269 A; 568 C; 491 G; 265 T;  
Query Match 1.38; Score 15; DB 23; Length 1593;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 395 agagaagctcagg 409  
|||||  
CP 1015 AGGAGAGCTTCAGG 1001

## RESULT 255

ID 850731 standard: DNA; 1599 BP.  
AC 850731-1997  
AC 850733;

DE 18-JUN-1990 (first entry)  
DE Genomic DNA from Plasmodium falciparum K1 isolates encoding the 41 MD  
DE merozoite antigen  
DE Antimalarial vaccine  
KW Antimalarial vaccine; positive immunisation; malaria diagnosis.  
FH Plasmodium falciparum.  
FS Key Location/Qualifiers  
FT cds 346..1434  
FT 619..622  
FT conflict  
FT /tag= b  
FT /note= "Differs from cDNA from M25 isolate of  
FT P. falciparum."  
FT /tag= c  
FT /note= "Differs from cDNA from M25 isolate of  
FT P. falciparum."  
FT 949..952  
FT conflict  
FT /note= "Differs from cDNA from M25 isolate of  
FT P. falciparum."  
PN PP-309746-A.  
PD 27-NOV-1988; 1140016.  
PR 08-SEP-1987; CR-003486.  
PI (Hoff) Hoffmann-La Roche AG.  
PI Certi U; 101095/14.  
DR P-PSDB: 893566.  
DE New peptide(s).  
PT contg. epitope(s) of Plasmodium falciparum merozoite antigen,  
PT for use in malaria vaccines  
PT Figure 1: Amino acid sequence of the 41 MD merozoite antigen  
CC The genomic DNA of P. falciparum K1 isolates was largely identical to this  
CC genomic isolate of K1, whereas cDNA from M25 isolates differs  
CC in 3 codons (see 82937). Plasmodium falciparum merozoite antigen can be  
CC coupled to an affinity peptide, or adsorbed or covalently  
CC cross-linked to a solid support. The peptide, as well as immunogens in  
CC anti-malaria vaccine. Antibodies directed against it are useful for  
CC passive immunisation and diagnosis.  
SQ Sequence 1599 BP; 619 A; 207 C; 259 G; 514 T;

Query Match 1.38; Score 14; DB 1; Length 1599;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1144 ggtgtgtatttt 1157  
|||||  
CP 546 GGTGTGTATTITTT 533

RESULT 256  
ID 062130 standard: cDNA; 1601 BP.  
AC 062130-1994 (first entry)  
DE 22-NOV-1994 (first entry)  
DE Endothelial specific receptor tyrosine kinase; tek gene; mouse;  
KW murine; chromosome 4; cardiotgenesis; ss.  
OS Mus musculus.  
PB Key Location/Qualifiers  
FT CDS 1..1601  
FT /tag= a  
FT /note= "open reading frame - does not begin with  
FT initiation codon. Thought to encode the  
FT cytoplasmic portion of a transmembrane RTK  
FT which is cleaved by a short C-terminal tail (33 amino acids)."  
FT 3'utr 906..1601  
FT /tag= b  
FT 03085201-A.  
PD 31-JAN-1994.  
PF 14-DEC-1992; 085291.  
PR 30-JUL-1992; US-921795.  
PI (MOON) MOON SINGH HOSPITAL CORP.  
PI (MOON) MOON SINGH HOSPITAL CORP.  
DR WP1: 94-126938/14.  
DR P-PSDB: 853147.  
DE Receptor tyrosine Kinase isolated from cells of the endothelial  
DE lineage - used for identifying (antagonists of ligand-receptor  
DE interaction).  
PS Example 1: Fig 2: 51pp; English.  
CC 062130 shows the nucleotide sequence of a 1.6kb tek cDNA isolated  
CC from a 13.5 day mouse embryo cDNA library. Translation of this  
CC sequence revealed a single large ORF that terminates with a TAG  
CC stop codon. The deduced amino acid sequence suggests that the several  
CC features of the deduced amino acid sequence suggest that the several  
CC tek cDNA encodes the cytoplasmic domain of a transmembrane RTK, the  
CC consisting of the catalytic domain followed by a C-terminal tail.  
SQ Sequence 1601 BP; 463 A; 335 C; 392 G; 411 T;  
Query Match 1.38; Score 14; DB 11; Length 1601;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 216 agcgtattttacc 229  
|||||  
QY 108 AGCGTATTTCAC 121

RESULT 257  
ID V04252 standard: cDNA; 1604 BP.  
AC V04252;



DT 05-JUN-1998 (first entry)  
 KW wild tomato acyltransferase cDNA.  
 OS glucose acyl ester; pest repellent; emulsifier; emollient; ds.  
 FH Lycopodium pennellii.  
 FH Key Location/Qualifiers  
 FT 1-1395  
 FT /tag= a  
 FT /product= acyltransferase  
 FT W09748911-A1.  
 PD 24-DEC-1997: H11005  
 PD 21-JUN-1996: US-655966  
 PR (CORR ) CORNELL RES FOUND INC.  
 PI Changas GS, Steffens JC;  
 DR WPI: 98-06119/06.  
 DT New acyltransferase from a wild tomato species - used to produce  
 PT glucose palmitoyl esters for use as pest repellents, emulsifiers and  
 PT emollients.  
 CC Claim 9: Pages 10-24; 34pp; English. tomato acyltransferase, which  
 CC can be used to produce glucose palmitoyl ester by reacting  
 CC 1-0-palmitoyl-beta-D-glucose with glucose or its partial palmitoyl  
 CC esters. Disclosed is a similar process for glucose acyl esters in  
 CC general. Keyated sugars are useful as pest repellents in  
 CC cosmetics, emulsifiers in foods or cosmetics and emollients in  
 CC cosmetics.  
 SQ Sequence 1604 BP; 515 A; 298 C; 325 G; 466 T;  
 Query Match 1.3%; Score 15; DB 39; Length 1604;  
 Best Local Similarity 100.0%; Pred. No. 347e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 775 gattccgagagta 789  
 Cp 963 GATCTCGAGAGTA 949

RESULT 258  
 AC Q00543.  
 DT 17-JUL-1990 (first entry)  
 DE Vampire bat plasminogen activator gene.  
 KW tissue plasminogen activator; vampire bat; Bat-PA(1); Bat-PA(H);  
 OS Desmodus rotundus.  
 FH Key Location/Qualifiers  
 FT 91-1524  
 FT /tag= a  
 FT /note= Claim 11.

FT sig\_peptide 91-198  
 FT /tag= b  
 FT mat\_peptide 199-1524  
 FT /tag= b  
 FT /product= Bat-PA(H)  
 FT /note= Claim 12.  
 PD EP-352119-A.  
 PD 24-JUN-1990: 307411  
 PD 20-JUN-1988: US-377221.  
 PR (MERRI ) Merck and Co.  
 PI Duong LT, Jacobs JW, Friedman PA, Dixon RAF, Gardell BJ, Mark GE,  
 DR WPI: 90-02461/04.  
 DR P-PSDB: R05122/23.25.  
 DT Vampire bat glycosylated plasminogen activating protein - which needs  
 PT fibrin cofactor to activate plasminogen, has greater selectivity for  
 PT human than plasminogen than PA.  
 CC The sequence is that of the longest cDNA clone isolated from a library  
 CC glands of vampire bats. An additional 800 nt of 3' UT sequence ending  
 CC in a poly(A) tail was also identified. The cDNA encodes a 336 amino acid  
 CC activator which, unlike tPA, does not contain the Kringle 2 domain and  
 CC plasmin-sensitive processing site. Three distinct species of activator  
 CC have been isolated corresponding to full-length (tag ), finger-minus,  
 CC and finger domain forms of tPA. These are designated Bat-PA(H),  
 CC Bat-PA(F), and Bat-PA(M). Bat-PA(H) is the most active. The new  
 CC proteins have greater selectivity towards fibrin-bound PA.  
 SQ Sequence 1620 BP; 434 A; 410 C; 440 G; 336 T;  
 Query Match 1.3%; Score 15; DB 1; Length 1620;  
 Best Local Similarity 100.0%; Pred. No. 347e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 649 tonaacattctgtgc 663  
 Cp 686 TCAAACTCTGCTGC 700

RESULT 359  
 AC Q49255.  
 DT 28-APR-1994 (first entry)  
 DE cad-4.  
 KW Long-distance homology; evolution; nematode;  
 OS Caenorhabditis elegans.  
 FH Key Location/Qualifiers  
 FT 91-1524  
 FT /tag= a  
 FT /note= Claim 11.

FR Key Location/Qualifiers  
 FT cda 1-1638  
 FT /tag= a  
 FT /product= cad-4\_gene\_product  
 PD W09320237-A.  
 PD 14-OCT-1993.  
 PF 01-APR-1993: D03102.  
 PF 01-APR-1992: US-861458.  
 PR (MERRI ) MERCK & CO INC.  
 PI Johnson CD, Marchionni MA.  
 DR WPI: 93-33693/42.  
 DR P-PSDB: R42742.  
 DT Long distance homology cloning of genes from lower organisms -  
 PT long distance homology cloning of genes from lower organisms -  
 PT amino acid sequences  
 PS Disclosure: Fig 9; 18pp; English.  
 CC The primers/probes (Q49266-Q49285) are used to isolate the cad-4  
 CC from the nematode C. briggsae.  
 SQ Sequence 1638 BP; 492 A; 351 C; 343 G; 456 T;  
 Query Match 1.3%; Score 14; DB 9; Length 1638;  
 Best Local Similarity 100.0%; Pred. No. 347e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 806 tatcaccaccaga 819  
 Cp 181 TATCACCACCAGA 168

RESULT 260  
 ID T80102 standard; cDNA; 1641 BP.  
 CC T80102 standard; cDNA; 1641 BP.  
 DE KF-1 gene open reading frame.  
 KW Alzheimer's disease; KF-1; human; mouse; KF361 gene; antibody; aa.  
 OS Homo sapiens.  
 PD 19-AUG-1997.  
 PF 20-JUN-1996: 181514.  
 PF 07-DEC-1995: JP-345659.  
 PR WPI: 95-024905/04.  
 DR P-PSDB: W26164.  
 DT A brain-specific expression gene - used in the diagnosis of  
 PT Alzheimer's disease  
 CC This gene encodes a protein that codes for evolutionary conserved  
 CC sequences shown in T80098-T80100 all represent cDNAs of the  
 CC invention, and are all KF-1 genes. The genes contain the open reading  
 CC frames represented by this sequence, T80101 and T80103. These genes,  
 CC their products and antibodies specific for the proteins are useful for

CC the diagnosis of Alzheimer's disease.  
 SQ Sequence 1641 BP; 489 A; 305 C; 351 G; 495 T;  
 Query Match 1.3%; Score 14; DB 34; Length 1641;  
 Best Local Similarity 100.0%; Pred. No. 347e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1229 gtgagagagagta 1242  
 Cp 982 GTGAGAGAGTA 995

RESULT 261  
 ID Q80315 standard; DNA; 1641 BP.  
 CC Q80315 standard; DNA; 1641 BP.  
 DE Human 38kDa FK-506 binding protein DNA.  
 KW Human 38kDa FK-506 binding protein; allograft rejection;  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT cda 141-1208  
 FT /tag= a  
 FT G0270340-A.  
 PD 04-JAN-1995.  
 PF 15-JUN-1994: 012000.  
 PF 25-JUN-1993: US-083944.  
 PR (MERRI ) MERCK & CO INC.  
 PI Smith J, Merck & Co Inc.  
 DR WPI: 95-024905/04.  
 DR P-PSDB: R64226.  
 DT New human 38 kD FK-506 binding protein and DNA encoding it -  
 PT useful for identifying cpts, able to block allograft rejection  
 PT in human and murine systems.  
 CC Claim 2: Pages 33-34; 44pp; English.  
 CC Q80315 encodes R64226 the human 38kDa FK-506 binding protein. The  
 CC protein can be used to identify cpts, with immunosuppressive  
 CC activity, able to block allograft rejection by preventing T-cell  
 CC activation. The protein can be used to identify cpts, with  
 CC immunosuppressive activity, able to block allograft rejection by  
 CC identifying, quantifying and purifying FK-506 and FK-506 like cpts.  
 SQ Sequence 1641 BP; 324 A; 550 C; 507 G; 260 T;  
 Query Match 1.3%; Score 14; DB 13; Length 1641;  
 Best Local Similarity 100.0%; Pred. No. 347e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 828 cactacggcagc 841  
 Cp 452 CACTACGGCAGC 465



```

RESULT 262 ID Q14835 standard; cDNA: 1642 BP.
AC Q14835 AC Q14835 (first entry)
AD 17-FEB-1992 AD 17-FEB-1992 encoding gene 8-30-84
AE 5-HT (ID): Parkinson's disease; migraine; anxiety; eating disorder:
EM K-M G-protein; 5-hydroxytryptamine; as.
OS Homo sapiens. Location/Qualifiers
FT Key 2681464
FT Key 2681464 /tag= A
FT W081174-A.
PD 14-NOV-1991 PD 14-NOV-1991 U03200.
PR 08-MAY-1990 PR 08-MAY-1990 US-520716.
PA (NEUR-) NEUROGENETIC CORP.
PI Weinshank RL, BARGHEK T, Hartig PR;
PS P-SPB1-211117/48.
DB Nucleic acid encoding 5HT-ID receptors and their antibodies -
PT used to treat and diagnose conditions caused by abnormal 5HT-ID
PT receptor expression e.g. dementia
CC There are two subtypes of 5HT-ID receptors encoded by gene 8-30-84
CC and gene 11, respectively. A human placental genomic library was
CC screened with a 1.3kb fragment of the dog clone npc-4 to identify 5
CC clones. Three clones collectively span the entire coding region of
CC gene 8-30-84. The longest clone encodes a protein identical to that
CC isolated from a human hippocampus library. A plasmid containing the
CC 8-30-84 sequence has been designated pcZEV-8-30-84 and has ATCC No.
CC 40790. The invention also covers antisense oligonucleotides able to
CC modulate the function of the 5HT-ID receptor and thereby prevent or
CC Parkinson's Disease, dementia, eating disorders, pathological
CC anxiety and migraine.
CC See Q4836 for gene 11.
SQ Sequence 1642 BP; 355 A; 492 C; 378 G; 417 T;
Query Match 1.3%; Score 14; DB 3; Length 1642;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 315 tttcccaggagacc 328
CP 164 TTTCCCGAGGACC 151

RESULT 263 ID T91738 standard; DNA: 1650 BP.
AC T91738 AC T91738 (first entry)
AD 20-JUN-1998 AD 20-JUN-1998 (first entry)
AE SC virus DNA fragment
DE SC virus DNA fragment

```

```

JC virus, agnoprotein, VP23, VP231, VP3, VP1, contour protein,
progressive multifocal leukoencephalopathy, aa.
JC virus.
OS 224658-A.
JC virus.
PN 10921-1996 (first entry)
PN 23-FEB-1996; 062167.
PR 23-FEB-1996; JP-062167.
PR (A1SE) 1 A1SIN SEIKI KK.
PR WP1: 97-484094/43.
PR WP2: 97-484094/43.
PR WP3: 97-484094/43.
PR WP4: 97-484094/43.
PR WP5: 97-484094/43.
PR WP6: 97-484094/43.
PR WP7: 97-484094/43.
PR WP8: 97-484094/43.
PR WP9: 97-484094/43.
PR WP10: 97-484094/43.
PR WP11: 97-484094/43.
PR WP12: 97-484094/43.
PR WP13: 97-484094/43.
PR WP14: 97-484094/43.
PR WP15: 97-484094/43.
PR WP16: 97-484094/43.
PR WP17: 97-484094/43.
PR WP18: 97-484094/43.
PR WP19: 97-484094/43.
PR WP20: 97-484094/43.
PR WP21: 97-484094/43.
PR WP22: 97-484094/43.
PR WP23: 97-484094/43.
PR WP24: 97-484094/43.
PR WP25: 97-484094/43.
PR WP26: 97-484094/43.
PR WP27: 97-484094/43.
PR WP28: 97-484094/43.
PR WP29: 97-484094/43.
PR WP30: 97-484094/43.
PR WP31: 97-484094/43.
PR WP32: 97-484094/43.
PR WP33: 97-484094/43.
PR WP34: 97-484094/43.
PR WP35: 97-484094/43.
PR WP36: 97-484094/43.
PR WP37: 97-484094/43.
PR WP38: 97-484094/43.
PR WP39: 97-484094/43.
PR WP40: 97-484094/43.
PR WP41: 97-484094/43.
PR WP42: 97-484094/43.
PR WP43: 97-484094/43.
PR WP44: 97-484094/43.
PR WP45: 97-484094/43.
PR WP46: 97-484094/43.
PR WP47: 97-484094/43.
PR WP48: 97-484094/43.
PR WP49: 97-484094/43.
PR WP50: 97-484094/43.
PR WP51: 97-484094/43.
PR WP52: 97-484094/43.
PR WP53: 97-484094/43.
PR WP54: 97-484094/43.
PR WP55: 97-484094/43.
PR WP56: 97-484094/43.
PR WP57: 97-484094/43.
PR WP58: 97-484094/43.
PR WP59: 97-484094/43.
PR WP60: 97-484094/43.
PR WP61: 97-484094/43.
PR WP62: 97-484094/43.
PR WP63: 97-484094/43.
PR WP64: 97-484094/43.
PR WP65: 97-484094/43.
PR WP66: 97-484094/43.
PR WP67: 97-484094/43.
PR WP68: 97-484094/43.
PR WP69: 97-484094/43.
PR WP70: 97-484094/43.
PR WP71: 97-484094/43.
PR WP72: 97-484094/43.
PR WP73: 97-484094/43.
PR WP74: 97-484094/43.
PR WP75: 97-484094/43.
PR WP76: 97-484094/43.
PR WP77: 97-484094/43.
PR WP78: 97-484094/43.
PR WP79: 97-484094/43.
PR WP80: 97-484094/43.
PR WP81: 97-484094/43.
PR WP82: 97-484094/43.
PR WP83: 97-484094/43.
PR WP84: 97-484094/43.
PR WP85: 97-484094/43.
PR WP86: 97-484094/43.
PR WP87: 97-484094/43.
PR WP88: 97-484094/43.
PR WP89: 97-484094/43.
PR WP90: 97-484094/43.
PR WP91: 97-484094/43.
PR WP92: 97-484094/43.
PR WP93: 97-484094/43.
PR WP94: 97-484094/43.
PR WP95: 97-484094/43.
PR WP96: 97-484094/43.
PR WP97: 97-484094/43.
PR WP98: 97-484094/43.
PR WP99: 97-484094/43.
PR WP100: 97-484094/43.
PR WP101: 97-484094/43.
PR WP102: 97-484094/43.
PR WP103: 97-484094/43.
PR WP104: 97-484094/43.
PR WP105: 97-484094/43.
PR WP106: 97-484094/43.
PR WP107: 97-484094/43.
PR WP108: 97-484094/43.
PR WP109: 97-484094/43.
PR WP110: 97-484094/43.
PR WP111: 97-484094/43.
PR WP112: 97-484094/43.
PR WP113: 97-484094/43.
PR WP114: 97-484094/43.
PR WP115: 97-484094/43.
PR WP116: 97-484094/43.
PR WP117: 97-484094/43.
PR WP118: 97-484094/43.
PR WP119: 97-484094/43.
PR WP120: 97-484094/43.
PR WP121: 97-484094/43.
PR WP122: 97-484094/43.
PR WP123: 97-484094/43.
PR WP124: 97-484094/43.
PR WP125: 97-484094/43.
PR WP126: 97-484094/43.
PR WP127: 97-484094/43.
PR WP128: 97-484094/43.
PR WP129: 97-484094/43.
PR WP130: 97-484094/43.
PR WP131: 97-484094/43.
PR WP132: 97-484094/43.
PR WP133: 97-484094/43.
PR WP134: 97-484094/43.
PR WP135: 97-484094/43.
PR WP136: 97-484094/43.
PR WP137: 97-484094/43.
PR WP138: 97-484094/43.
PR WP139: 97-484094/43.
PR WP140: 97-484094/43.
PR WP141: 97-484094/43.
PR WP142: 97-484094/43.
PR WP143: 97-484094/43.
PR WP144: 97-484094/43.
PR WP145: 97-484094/43.
PR WP146: 97-484094/43.
PR WP147: 97-484094/43.
PR WP148: 97-484094/43.
PR WP149: 97-484094/43.
PR WP150: 97-484094/43.
PR WP151: 97-484094/43.
PR WP152: 97-484094/43.
PR WP153: 97-484094/43.
PR WP154: 97-484094/43.
PR WP155: 97-484094/43.
PR WP156: 97-484094/43.
PR WP157: 97-484094/43.
PR WP158: 97-484094/43.
PR WP159: 97-484094/43.
PR WP160: 97-484094/43.
PR WP161: 97-484094/43.
PR WP162: 97-484094/43.
PR WP163: 97-484094/43.
PR WP164: 97-484094/43.
PR WP165: 97-484094/43.
PR WP166: 97-484094/43.
PR WP167: 97-484094/43.
PR WP168: 97-484094/43.
PR WP169: 97-484094/43.
PR WP170: 97-484094/43.
PR WP171: 97-484094/43.
PR WP172: 97-484094/43.
PR WP173: 97-484094/43.
PR WP174: 97-484094/43.
PR WP175: 97-484094/43.
PR WP176: 97-484094/43.
PR WP177: 97-484094/43.
PR WP178: 97-484094/43.
PR WP179: 97-484094/43.
PR WP180: 97-484094/43.
PR WP181: 97-484094/43.
PR WP182: 97-484094/43.
PR WP183: 97-484094/43.
PR WP184: 97-484094/43.
PR WP185: 97-484094/43.
PR WP186: 97-484094/43.
PR WP187: 97-484094/43.
PR WP188: 97-484094/43.
PR WP189: 97-484094/43.
PR WP190: 97-484094/43.
PR WP191: 97-484094/43.
PR WP192: 97-484094/43.
PR WP193: 97-484094/43.
PR WP194: 97-4
```

PR 13-SEP-1994: JP-218392.  
PR 07-DEC-1994: JP-303809.  
PA (SANY) SANKYO CO LTD.  
PI Kawahara, 118, Koishi R.  
PI 19860608 492030.  
P P-5008 492030.  
PT Clover yellow vein virus nuclear inclusion and di-chloro:indomethol  
PT oxidized glutathione reducing protein - useful in autolyzing  
PT fusion protein expression system and for treating diseases related  
PT to glutathione deficiency, reduced oxygen, respectively.  
PS Claim 47: Page 134-137; 168pp: English.  
PS C A CDNA clone (165092), designated 31-7, codes for a human  
PS glutathione reducing protein, designated RMI-7 (492050).  
PS The RMI-7 cDNA encodes a protein of 140 amino acids.  
PS The CDNA has been inserted into a vector to allow  
PS expression in *Escherichia coli* cells. Mature RMI-7 shows reducing  
PS activity toward dichloroindemol and oxidized glutathione, and  
PS caused by activated oxygen species related to oxidative stress or  
PS sequenced 488 A: 321 C: 433 G: 408 T:  
SQ Sequence 1650 BP:  
SQ Query Match Score 14: BP 37: Length 1650:  
SQ Query Local Similarity 100.0%: Prod No: 347+002: 1650:  
SQ Matches 14: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
Db 1447 gncatgcagcgt 1460  
Cd 687 GACATGACGCGT 674

```

DNU m3], encoding human nuclear factors of activated T cells...
useful for screening potential therapeutic and diagnostic agents for
immune system diseases
PT Disclosures, Page 55-56, 64pp; English.
PT cDNA sequence (J36971), comprises the 3' end of human nuclear
PT factor 4, NFAT4, also known as NFATp, NFATp4, NFATp4a, NFATp4b,
PT Accession No. U00001, Genbank, EMBL, DDBJ, NCBI, Nucleotide
PT Accession No. 700-1068 (M02351). 2 Other classes of NFAT4 cDNAs, NFAT4a
PT (J33677) and NFAT4b (J35975), were also isolated; the NFAT4c
PT transcript was most abundant. The major sites for NFAT4 expression
PT were identified in the spleen, thymus, and peripheral blood lymphocytes
PT and were isolated from cDNA libraries prep'd. from Jurkat T-cells and
PT human peripheral blood lymphocytes by screening with a probe derived
PT from the left similarity region of the human NFATp gene. NFAT genes
PT encode a family of transcription factors, as probes and primers;
PT and (see abstract) to modulate cellular expression of intracellular
PT availability of active NFAT.
SQ Sequence 1662 BP: 418 C: 354 G: 489 T:
Query Match 1.13; Score 14; DB: 23; Length 1662.
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1447 attcgtcttttttata 1460
|||||ttttttttttttttt
QY 163 CTCTGCTTTTATA 196
RESULT 266

```

RESULT	265
TM	standard: cDNA; 1662 BP.
AC	T18971
CC	T36871
DT	18-NOV-1996 (first entry)
DE	Human transcription factor NFATc4 (3' end) cDNA.
NW	Nuclear factor of activated T-cells; NFAT; NFATc4;
KW	transcription factor; binding assay; immune system disease;
TX	reverse transcriptase-polymerase chain reaction; RT-PCR;
OS	Homo sapiens.
FT	Key
FH	Location/Qualifiers
FI	1..1110
FT	cds
PI	/tag= a
PH	W069626959-11
PD	06-SEP-1996
PF	04-MAR-1996; UQ3113.
PR	02-MAR-1995; US-356479.
PA	(TULIA-) TULAKIR INC.
DP	USPTO
HPY	06-412728/41.
DR	P-PSDB: W02253.
DR	P-PSDB: W02253.

[illegible]















```

Beat Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;

db 247 ttgacttcagctga 260
      6 ttgcacttcagctga 19

RESULT 270
ID.DD 789220 standard; DWA, 1764 BP.
AC AC
DE 28-APR-1998 (first entry)
DE Streptococcus pneumoniae aspartyl tRNA synthetase gene.
DE This gene encodes the aspartyl tRNA synthetase, an enzyme
DE involved in the synthesis of tRNA. The gene is located on
DE vaccine gene therapy. Immune response: meningitis, otitis media;
DE conjunctivitis; pneumonia; bacteremia; ss.
DE Streptococcus pneumoniae.
DE Streptococcus pneumoniae.
DE Location/Qualifiers
DE     cds            Location/Qualifiers
DE     CDS             /tag= a
DE     FT              /product= 'asps'
W93739014.AL.
CDS             1..1764
FT              16-APR-1997; U06638.
FT              16-APR-1997; GB-007993.
FT              16-APR-1998; GB-007993.
FT              (SHK ) SKITHKLNE BEECHAN CORP.
FT              (SHK ) SKITHKLNE BEECHAN PLC.
FT              16-APR-1998; GB-007993.
FT              WPI: 97-526390/48.
DR DR
DR P-SDS: W139156.
DR encoding aspartyl-tRNA synthetase from Streptococcus pneumoniae
DR encoding aspartyl-tRNA synthetase from bacterial infections
DR Claim 4, Page 33; 45pp; English.
DR The present sequence encodes aspartyl tRNA synthetase (asps) from
DR Streptococcus pneumoniae. A host cell containing a vector comprising
DR Streptococcus pneumoniae. The asps gene, asps, is expressed to
DR polypeptide encoding asps, asps through the asps gene. The
DR induce an immunological response in a mammal to generate antibodies
DR and/or t cell immune response to protect against disease. The antibodies
DR protect against invasion of bacteria, e.g. by blocking adherence of
DR bacteria to mucous membranes. The antibodies also protect against
DR tissue caused by mechanical, chemical or thermal damage of by
DR implantation of in-gelling devices, or wounds in the mucous membranes.
DR Antagonists are used to inhibit the asps protein.
DR Sequence 1764 BP; 491 A; 363 C; 434 G; 476 T;
DR Query Match 1.3%; Score 14; DB 37; Length 1764;
Beat Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0;

db 853 ttgcacttcagctga 966

```

775 TTTTCAAGCGAGA 762

CCP

RESULTS 280  
IZ Q14340 standard; DNA; 1767 BP.  
AC Q14340;  
17-JAN-1992 (first entry)  
DE Mouse epithelin precursor  
DE Mouse epithelin precursor  
DE Mus musculinus  
FY Key Location/Qualifiers  
FH Key Location/Qualifiers  
FT 8.11774  
TT /tag= "clain 29, page 56"  
TT /note="clain 29, page 56"  
PD W09115510-A.  
PN 17-OCT-1991.  
PP 03-APR-1991. U02321  
PP 03-APR-1991. U02321  
PR 13-MAR-1991. U084706  
PR (BRIN ) BRISTOL-MYERS SQUIB.  
PA Shoyab M, PLOWMAN GD;  
PI Shoyab M, PLOWMAN GD;  
PR P-5250. R14337/44.  
DR P-5250. R14337/44.  
DT New cysine-rich growth modulating protein, epithelins - useful  
as inhibitors of neoplastic cell growth and to promote wound  
healing and treat psoriasis.  
DT Ep-1 and Ep-2 were isolated from rat kidneys and their amino acid  
sequences determined. A full length rat Ep-2 cDNA (Q14338) was obtd.  
by screening a rat kidney cDNA library in lambda gt10 with PCR  
amplified probes. These probes were also used to obtain the mouse  
Ep-2 cDNA (Q14340) from a mouse kidney cDNA library.  
CC An anti-sense ribonucleic acid molecule complementary to the  
indicated fragment in the features is also claimed.  
CC The growth of some cell types while inhibiting the growth of others  
Ep-2 is functionally similar to Ep-1 w.r.t. growth inhibitory  
bioactivity. In contrast, however, Ep-2 is apparently not capable of  
affecting the growth stimulatory activity characteristic of Ep-1 and, see also Q14338-40, Q14352-53, R14338-9 and R15315-20.  
CC See also Q14338-40, Q14352-53, R14338-9 and R15315-20.  
CC The growth 1767 BP; 360 G; 492 G; 412 T;  
CC QUARY MATCH 130; SCORE 14; DB 3; Length 1767;  
CC Best Local Similarity 100.00; Ref. No. 3.47e+02; 1767;  
CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps  
DB 726 cttctgctgctgcgc 739

891 CCGCAGCTGCTGCC 904

RESULT 281  
AC Q14336: standard; cDNA, 1767 BP.  
AC Q14338:  
DT 17-JAN-1992 (first entry)  
DB rat epithelin precursor.  
DS anti-sense RNA: growth regulation; inhibition; stimulation; ss.  
SS Ratian status.  
FE Key Location/Qualifiers  
FH misc\_rna 31..1797  
FT /\*seq= "Clain 19, page 55"  
FN /note=  
W09115510-A.  
PD 17-OCT-1991.  
PP 03-APR-1991: D02321.08  
PR 13-MAR-1991: US-081796.  
PA (BRIM ) BRISTOL-MYERS SQUIB.  
PI Shoyab M, Plovean CD;  
PM Wintig JI 32:158/44.  
PN  
PT New cytosine-rich growth modulating proteins, epithelins - useful as inhibitors of neoplastic cell growth and to promote wound healing and treat psoriasis  
CC The ER-2 gene has been cloned from rat kidneys and their amino acid sequence determined. It is 1879 bp long.  
CC Sequences determined. A full length rat ER-2 cDNA was obtained by screening a rat kidney cDNA library in lambda gt10 with PCR generated ET probes.  
CC These probes were also used to obtain the mouse Et gene (Q14340) from a human genomic DNA library. ET DNA was also obtained from human sources (Q14339).  
CC An anti-sense ribonucleic acid molecule complementary to the indicated fragment in the features is also claimed.  
CC ER-2 is a functional growth regulator capable of stimulating growth in primary cultures of epithelial cells and fibroblasts.  
CC ER-2 is functionally similar to ET-1 w.r.t. growth inhibitory bioactivity. In contrast, however, ET-2 is apparently not capable of eliciting the growth stimulatory activity characteristic of ET-1 and, therefore, may have different biological functions.  
CC See also Q14338-Q14952-53; R14328-9 and R15315-20.  
SQ Sequence 1767 BP; 359 A; 499 C; 491 G; 418 T;  
Query Match 13%; Score 14; DB 3; Length 1767;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 726 ccgcagctgctgcc 739  
891 |111111111111| 904

```

RESULT 282
ID Q14339 standard: DNA; 1779 BP.
ID Q14339 standard: DNA; 1779 BP.
D1 11438; 1992 (first entry)
D2 Human epithelin precursor.
D3 ER: anti-sense RNA; growth regulation; inhibition; stimulation; as.
KW Homo sapiens. Location/Qualifiers
FT misc_jna
FT 41..1819
FT /*tag= a
FT /*note= *Claim 9, page 54*

W0115510.A.
1779-CTC-1991.
PP 03-APR-1991: 501321.
PP 03-APR-1990: US-004508.
PP 13-MAR-1991: US-083796.
PR Shoyab M, Ploewman GD.
PR WPI: 91-325168/44.
PR P-PDB: R14336.
DR New cytoskeletal growth modulating proteins, epithelins - useful
DR in promoting cell growth and to promote wound
DR healing and treat psoriasis [see also:
DR Disclosure: Fig 22: 97pp: English.
PS ER-1 and ER-2 were isolated from rat kidneys and their amino acid
PS sequence determined. ER-1 is identical with cDNA (Q14339) and was
PS generated by screening a rat kidney cDNA library with ER-1 probe.
PS generated ER probes. These probes were also used to obtain the mouse
PS ER gene (Q14340) from a mouse T-cell genomic library.
PS ER DNA was also obtained from human sources (Q14339).
PS ER-1 is a functional growth regulator, capable of stimulating
PS the growth of some cell types while inhibiting the growth of others.
PS ER-1 functions in contrast to ER-2. ER-2 is growth inhibitory
PS bioactive in contrast to ER-1. ER-1 is growth stimulatory
PS eliciting the growth stimulatory activity characteristic of ER-1 and,
PS in fact, antagonizes this ER-1 activity.
SC See also Q14338-40, Q19522-53, R14328-9 and R15315-20.
SC Sequence 1779 BP, 304 A, 576 C, 547 G, 350 T;
Query Match 1.3%; Score 14; DB 3; Length 1779;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps
DB 729 cctgcagctgctgc 742
DB 991 ccttcagctgcctc 904

```















```

PH Key Location/Qualifiers
FT signal_peptide 24-63
FT /*tag = a
FT mat_peptide 64..790
FT /*tag = b
PD WO9203470-A-
PD 05-MAR-1992
PF 28-ACG-1991; D06140
PF 28-ACG-1990; DS-574613
PF 28-ACG-1990; DS-574613
PF 28-ACG-1990; DS-574613
PA (CHIR)-CHIRON CORP.
PA Kiefer MC, Maslart F;
PI WPI; 92-066830/J2.
PR DNA DBSIB; 018686
PT DNA; human-like growth factor binding protein-5 - used
PT to obtain large quantities of protein useful in e.g. wound
PT healing
PT Clin 2; pag 1; 66pp; English.
CC fragment, having at least 10 nucleotide DNK seq. encoding KGP BP-5 or a
CC fragment, having greater than or equal to 10 nucleotide DNK seq. encoding
CC BP-5 as pref. of human origin and is genomic or cDNA pref. contained
CC in pBSP-5.
CC In pBSP-5.
SQ Sequence 1938 BP: 368 A; 576 C; 584 G; 410 T;
Query Match 1.39; Score 14; DB 3; Length 1938;
Best Local Similarity 100.0%; Pred.No. 3,47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 540 tctcttcagacgca 553
| | | | | | | | | |
QY 832 tctcttcagacgca 845
RESULTS 297
ID T95211 standard; DNA; 1961 BP.
QY T95211; 1998
DZ Human disintegrin metalloprotease (ecogrocinase) gene.
KW Diintegrin; metalloprotease; arthritis; osteoarthritis; diagnosis;
KW aggrecanase; arthropathy; osteoporosis; anticytokine spondylitis;
KW Bone tissue disease; human; ss.
FH Homo sapiens.
FV Key Location/Qualifiers
FT CD8 2..1477
FT WO921931-11
FT PD 04-SEP-1997
FT PD 28-FEB-1997; Q03217.
FF 01-MAR-1996; US-012679.
FF (USCA) PATENT OFFICE CO
FA (UYCA) UNIV CANE WESTERN RESERVE.

```

```

PR 19 DEC-1985   US-574695
PR (DAND) DANA FARBER CANCER INST INC
PI Young I., Shin J., Strominger JL, Veldamdi RK:
PI WPI: 97-341353/31.
PR PDBS: WJ1189.
PI 10/12/85.
PT treatment of autoimmune disease and for T and B cell proliferation,
PT e.g. for treatment of tumours
PS claim 4; Fig 3: 1'5pp: English.
PS 10/12/85.
CC 77.5 per cent homologous to the nuclear acid nucleoside p53 encoded
CC in 783133. This cytoplasmic polypeptide is expressed in B cells and
CC other cells of haematopoietic origin e.g. T cells. p53 is capable of
CC modulating or B cell development and/or for B cell activation e.g. by
CC binding to the B cell receptor. The polypeptide is involved in the
CC expression of cellular proteins e.g. cell cycle regulatory proteins attenuating cell
CC expression of cell cycle dependent kinase inhibitors and arresting cell
CC cycle progression at specific boundaries to thereby modulate cell
CC growth. The polypeptide has therapeutic uses in immune system related
CC treat disorders where this is beneficial e.g. infections by retrovi-
CC microorganisms, e.g. bacteria, viruses and protozoans. p53 can be used
CC to expand T cell populations for treating infectious diseases or cancer,
CC any infecting agents may be transduced to render them resistant to
CC responses and may be used to treat a variety of autoimmune diseases, e.g.
CC diabetes mellitus, arthritis, multiple sclerosis allergic reactions,
CC Crohn's diseases etc. 469 A. 531 C: 546 G: 431 T;
SQ Sequence 437 BP, 469 A. 531 C: 546 G: 431 T;

Query Match      1.3%; Score 14; DB 36; Length 1777;
Best Local Similarity 100.0%; Pred.No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      940 gsgagatgactgctg 953
Cc      510 GGAGATGACTCTG 497

RESULT 299
ID N90225 standard; DNA; 1982 BP.
DE 1-NOV-1989 (first entry)
DE Malaria-specific Plasmodium falciparum protein.
KW DNA; malaria; 41KD protein; Plasmodium falciparum; vaccine.
OS Plasmodium falciparum/Qualifiers
OS 526.1611
OS cda
PF */tag= a
PI EP-32712-A.
PI 20 DEC-1988.
PI 20 DEC-1988. 121289

```







Db	832	testcangattott	845
Query Match	138;	Score 14;	DB 24;
Best Local Similarity	100.00;	Pred. No	3.47e+02;
Matches	14;	Conservative	0;
		Mismatches	0;
		Indels	0;
Gap	0;		

PT 17-NOV-1993 011153  
PT 17-NOV-1993 02-3745Z.  
PA (ICGS) ICOS CORP.  
PA Godaka R., Gray PW, Schweickart VI;  
DR 94-200264/24.  
PT DNA encoding seven trans-membrane receptors - used to develop  
PT prods. for use as therapeutic or diagnostic agents for conditions  
PT involving the receptors. Orig. English.  
PT Two primers (G66148, G66150) used to amplify human genomic DNA  
CC purified from leukocytes. Approximately 1000 clones were isolated  
CC after the initial amplification reaction and probed with sequences  
CC specific for seven transmembrane receptors 158R1, ATRR and R20.  
CC Analysis of the clones revealed that approximately 10% of the  
CC clones were identical to the known sequences of the receptors.  
CC Seven new clones were identified that appeared to encode  
CC trans-membrane receptor segments. Two more primers (G66151,  
CC G66152) were used to isolate a full length version of one of these  
CC clones which was designated G66153.  
SQ Sequence 2058 BP; 472 N; 584 C; 448 T;  
Query Match 1.3%; Score 14; DB 11; Length 2058;  
RefSeq Local Similarity 100.0%; Pred. Ref: J47602;  
Accession ID: Conservative G; Mismatches 0; Indels 0; Gaps 0;

Dt 568 ggcagctactcttt 531  
Qt 244 cggacgaccatccttt 257

RESULT 308  
AC 2000 standard; cDNA; 2080 BP.  
AC TR0098  
DT 25-NOV-1997 (first entry)  
DE KF-1 gene.  
RW Alzheimer's disease; KF-1; human; mouse; KP361 gene; antibody; ss.  
KW Key word options. Location/Qualifiers  
cds 1..1641  
FT FT J09315495.A  
FT 18-AUG-1997  
PP 20-JUN-1996; 181514.  
PR 07-DIC-1995; JP-345659  
PR UNUMO SUGIYOMO SEITAKU KK.  
VF 07-DEC-1996  
P-PDB; M36164.  
PT A brain-specific expression gene - used in the diagnosis of  
PT Alzheimer's disease  
CC This sequence has 2 PDB: Japanese.







10kb were ligated into the BamHI site of pIL2352. One recombinant plasmid contained a 6.4 kb insert coding for the Abi.105 mechanism. This plasmid was designated pIL305. A 700bp EcoRV-XbaI fragment was excised from pIL305 and the religated plasmid, containing a 6.4 kb insert, was designated pIL438. The Abi.105 gene was partially sequenced and compared with the published sequence [19]. The sequence encodes the transposase from insertion sequence IS81. The sequence given here is that of ORF7, which codes for a protein involved in the bacteriophage resistance mechanism; part of the IS81 transposase gene and one of the inverted repeats (IR) sequences are boxed. See also pIL352 and Q23-716 (15) sequences. Sequence 243-794, 332-378, 602 T.

```

Query Match      1.3%      Score 14: DB 4: Length 2094:
Best Local Similarity 100.0%      Pred. No. 3.47e+02:
Matches 14: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 586 aaagaacacaaaga 599
|||||
Cp 645 AAGAAACACNAGA 632

```

	RESULT	313	ID T36141 standard; cDNA: 2096 BP.
	AC	T36141:	KO-870
	DE	Rumex L-saparinase (cist enzyme)	
	DE	Rumex L-saparinase cDNA	
	LK	L-asparaginase; antitumor; cancer;	lymphoma; leukemia; therapy;
	KW	ss.	
	PB	Homo sapiens.	
	PH	cds	Location/Qualifiers
	FT		93..1814 / *tag= a
	FT	BP-7760312.A2.	
	PR	14-AUG-1994.	
	PP	07-FEB-1996: 300803.	
	PR	08-FEB-1996: JP-042564.	
	CA	(HAY) HAYASHIDARA SEIBETSU KAKAO.	
	DN	WTI : 96-364333/30	N. Tanaka M., Torigoe K.
	DR		

DR p-FSDB; W02105.  
PT DNA encoding mammalian L-asparaginase - for prodn. of recombinant  
P1 enzyme. Genbank accession # F02105.  
P5 Cloning strategy. Genbank accession # F02105.  
CC A CDNA clone (736141) codes for human L-asparaginase (W02105). To  
CC obtain the cDNA, a human liver cDNA library was screened with guinea  
CC pig L-asparaginase cDNA (736138). The cDNA can be used with guinea  
CC pig recombinant L-asparaginase, an antitumor agent, in transformed  
CC host cells. Genbank accession # F02105.  
SQ Sequence. Genbank accession # F02105. 357 A; 653 C; 685 G; 361 T;

Tue Nov 17 08:55:26 1998

US-08-887-977-9.119

OS	Dendodus rotundus.				
PH	Kennedys	Location/Qualifiers			
FT	cds	180..1364			
FT	/stage	a			
FT	/product=w-P4_gamma				
PD	EP-383417-A.				
PD	22-Aug-1990.				
PP	13-FEB-1990:	250043.			
PP	13-FEB-1989:	DE-904580.			
PP	13-MAY-1989:	DE-904580.			
PB	SCHENKEL SCHENKEL AG.				
P1	Baldus B. Donner P.	Scheunung WD, Alagon A, Boifod W;			
P1	Kratzschmar JZ, Haendler BJ, Langer G,				
DR	RPT:	90-235697/34.			
DR	New fibrin specific thrombolytic agent v-PA - isolated from				
PT	saliva of Desmodus sp. bats.				
PT	Claim 5; Fig 26d: 49pp: English.				
PS	The sequence was obtained from gamma pp.				
CC	claim 5; Fig 26d: 49pp: English.				
CC	bats. The v-PA protein produced on expression of the gene is a				
CC	low molecular form consisting of a Kringle domain and a protease				
CC	domain. It is fibrin specific and dissolves blood clots.				
CC	See also EP030000657 for A.				
SI				512 C.	466 F.

```

Query Match      1.31;  Score 15;  DB 1:  Length 2100;
Best Local Similarity 100.0%;  Pred. No. 8.37e+01;
Matches 15;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Db      489  tcaaacctctgtgc 503
ov      686  tcaaacctctgtgc 700

```

	RESULT	316 9806	standard; DNA; 2107 BP.
AC	T9800:		
DE	20-MAR-1998	(first entry)	
DE	Human steroid Salpha-reductase coding sequence.		
KW	human; ss.		
KW	Steroid Salpha-reductase; enzyme; testosterone conversion; inhibitor;		
PN	Rom sapiens.		
PT	CDS	Location/Qualifiers	
PT	CD5	31..810	
PT	PT	/tag= a	
PN	US5679521-A.		
PD	21-OCT-1997.	457616	
PR	01-JUN-1995.	45765459	
PR	10-APR-1990.	US-517661	

Page 315

Page 216

THE NOV 17 08:55:26 1998

01-JUN-95 US-457616  
 (TEMA) UNIT VEGANS SYSTEM.  
 Anderson S, Russell DW.  
 WPI: 97-53718/48.  
 Pre-SDB: W4444.  
 The present invention relates to a recombinant steroid salpha-reductase enzyme - by culturing cell containing DNA encoding the enzyme.

**Disclosures:** Columns 45-50; 70pp; English.

This sequence represents the DNA encoding human steroid salpha-reductase. This method is for producing a steroid salpha-reductase enzyme. The method is for preparing a recombinant host cell containing a DNA segment encoding a steroid salpha-reductase and culturing the cell under conditions such that the steroid salpha-reductase is produced by the cell. The steroid salpha-reductase can be used as a reagent or as a component of a kit for assaying substances that affect the enzymatic activity of steroid salpha-reductase. Substances identified as inhibiting steroid salpha-reductase activity can be used for inhibiting the conversion of

Sequence 2107 BP: 475 A: 482 C: 486 G: 664 T

```

Query Match      1.3%      Score 14: DB 37: Length 2107;
Best Local Similarity 100.0%:   Pred.No. 3.47e+00;
Matches 14: Conservative      0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT	317
ID	G97380 standard; DNA; 2107 BP.
AC	G97380;
DT	01-APR-1996 (first entry)
KT	Human type I steroid 5-alpha reductase cDNA.
KW	Steroid 5-alpha reductase; sexual development; differentiation; Proteobiont; inhibit; prostatic hyperplasia; acne; hirsutism; androgens; androgenesis; prostate cancer; testosterone; dihydroxytestosterone; aa.
OR	Homo sapiens.
FE	Key
FT	Location/Qualifiers
FF	31..810
FD	/tag= a
PN	/Product= steroid_5_alpha_reductase_type-I
PD	US9542262-A.
PI	06-JUN-1995.
PF	30-APR-1990; PF 517661.
PP	30-APR-1990; US-517661.
PR	18-NOV-1991; US-795859.
PT	(CERVA) UNIV WISCONSIN.
PV	Audrey S.



DR WP1: 95-214659/28.  
 CC Steroid 5 alpha-reductase nucleic acid segments and recombinant  
 PT vectors, where the sequences are useful in e.g. analysis of normal  
 PT and abnormal sexual differentiation  
 PS Claim 3: Column 47-52; 7pp; English.  
 CC dihydroxyacetone. The human enzyme catalyzes the conversion of testosterone to  
 CC dihydroxyacetone. The human enzyme has been isolated and purified. The human enzyme  
 CC cDNA sequence (Q57180) has been isolated and purified. The human enzyme  
 CC cDNA can be used in the prep. of genetic constructs for the large  
 CC scale production of SHD5A or as probes for enzyme-encoding sequences  
 CC in cDNA libraries. The human enzyme is also useful in the analysis  
 CC of normal and abnormal sexual differentiation, hypogonadism, and  
 CC hyperplasia, male pattern baldness; acne; hirsutism and endometriosis.  
 SQ Sequence 2107 BP: 475 A; 482 G; 486 G; 664 T;  
 Query Match 1.3%; Score 14; DB 17; Length 2107;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 379 tctgactctgcgc 392  
 Cp 419 tctgactctgcgc 406

RESULT 318  
 ID Q14387 standard; DNA: 2107 BP.  
 AC Q14387;  
 DT 10-FEB-1992 (first entry)  
 DE Human steroid 5-alpha reductase gene.  
 KW Human steroid 5-alpha reductase; gene; hirsutism; dihydroxyacetone; ss.  
 OS Homo sapiens.  
 FN Key Location/Qualifiers  
 FT cds 31..810  
 FT /tag= a  
 FT /tag= b  
 FT polyA\_signal 2582..2597  
 FT W05117251-A.  
 PD 24-NOV-1991.  
 PD 30-APR-1990; US-517661.  
 PA (TEMA) UNIV TEXAS SYST.  
 PI Anderson S, Russell DW;  
 DR P-5058; R3140/2/48.  
 PT New DNA encoding human steroid 5 alpha-reductase - and corresp.  
 PT vectors and host cells, useful to screen potential inhibitors for  
 PT treatment of endocrine abnormalities, etc.  
 PS Claim; Fig 7; 8pp; English.  
 CC Rat cDNA probes (see Q14386) were used to screen a human prostate

CC cDNA library, to identify a 2.1 kb fragment encoding a steroid  
 CC 5-alpha reductase of 259 amino acids. The invention also covers  
 CC fragments of the human-derived sequence as well as considering  
 CC the human-derived sequence as well as considering  
 SQ Sequence 2107 BP: 475 A; 482 G; 486 G; 664 T;  
 Query Match 1.3%; Score 14; DB 3; Length 2107;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 379 tctgactctgcgc 392  
 Cp 419 tctgactctgcgc 406

RESULT 319  
 ID Q14387 standard; cDNA: 2109 BP.  
 AC Q14387;  
 DT 27-APR-1998 (first entry)  
 DE Human calpain cDNA.  
 KW Calpain; human; leukocyte; calcium dependent cysteine protease;  
 KW screening; activator; inhibitor; treatment; prevention; cancer;  
 KW Alzheimer's disease; myopathy; cataracts; collagen disease;  
 KW ischaemic heart disease; atherosclerosis; arthritis; da.  
 OS Homo sapiens.  
 FN Key Location/Qualifiers  
 FT cds 1..2109  
 FT /tag= a  
 FT /product= calpain  
 FT /note= "stop codon not given"  
 PD 08-OCT-1997.  
 PD 03-APR-1997; 105308.  
 PR 05-APR-1996; JP-083649.  
 PA (TAKE) TAKEDA CHEM IND LTD.  
 PI Takeuchi T, Shimada K, Shimada Y;  
 DR WPI: 97-482674/45.  
 DR P-5058; R41564.  
 PT Human calpain protein and related DNA - useful for drug screening  
 PT and treating cancer, stroke, etc.  
 PS Claim; 10pp; English.  
 CC The present sequence encodes calpain, a human leukocyte derived  
 CC calcium dependent cysteine protease. Calpain can be used to screen  
 CC for compounds that activate or inhibit its proteolytic activity.  
 CC Calpain DNA can be used to treat or prevent cancer, cerebral  
 CC Alzheimer's disease, myopathy, cataracts, ischaemic heart  
 CC disease, atherosclerosis, arthritis or collagen disease.  
 SQ Sequence 2109 BP: 499 A; 529 G; 623 G; 458 T;

Query Match 1.3%; Score 14; DB 37; Length 2109;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 797 tctgactctgcgc 810  
 Cp 19 tctgactctgcgc 6

RESULT 320  
 ID Q56794 standard; DNA: 2124 BP.  
 AC Q56794;  
 DT 27-APR-1998 (first entry)  
 DE Granulin; keratinocytes; wound healing; inhibition; peptide;  
 KW Granulin; keratinocytes; wound healing; inhibition; peptide;  
 KW granulocytes; leukocytes; ss.  
 OS Homo sapiens.  
 FN Key Location/Qualifiers  
 FT misc\_difference 1375..1381  
 FT /tag= a  
 FT /transl\_except= ATG encodes Valine.  
 FT /transl\_except= CAG encodes Glycine.  
 PD 05-APR-1993.  
 PD 03-FEB-1992; G08098.  
 PA (SOLJ) SOLOMON S.  
 PI Solomon S;  
 DR WPI: 93-32028/40.  
 PT New cysteine rich granulin peptide(s) from leukocyte(s) - are  
 PT keratinocyte inhibitors useful topically for wound healing  
 PS Disclosure: Figure 4c; 53pp; English.  
 CC The granulin inhibits keratinocytes and is useful in formulations  
 CC for the treatment of wound healing.  
 SQ Sequence 2124 BP: 385 A; 685 C; 630 G; 426 T;  
 Query Match 1.3%; Score 14; DB 9; Length 2124;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 754 cctgactctgcgc 767  
 QY 681 cctgactctgcgc 904

RESULT 321  
 ID Q6392 standard; cDNA: 2139 BP.  
 AC Q6392;  
 DT 03-APR-1993 (first entry)  
 DE Sequence of clone pVAM05R3 which contains probes for the in vitro  
 DE diagnosis of patients presenting an alteration on their chromosome  
 KW PCR primer: Charcot-Marie-Tooth type 1 disease; in vitro diagnosis;  
 KW chromosome 17p; ss.  
 OS Homo sapiens.

Query Match 1.3%; Score 14; DB 11; Length 2129;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 878 gctgactctgcgc 891  
 Cp 376 gctgactctgcgc 363

RESULT 322  
 ID Q0563 standard; DNA: 2133 BP.  
 AC Q0563;  
 DT 03-APR-1993 (first entry)  
 DE Sequence of clone pVAM05R3 which contains probes for the in vitro  
 DE diagnosis of patients presenting an alteration on their chromosome  
 KW PCR primer: Charcot-Marie-Tooth type 1 disease; in vitro diagnosis;  
 KW chromosome 17p; ss.  
 OS Homo sapiens.

Query Match 1.3%; Score 14; DB 11; Length 2129;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 878 gctgactctgcgc 891  
 Cp 376 gctgactctgcgc 363

RESULT 323  
 ID Q0563 standard; DNA: 2133 BP.  
 AC Q0563;  
 DT 03-APR-1993 (first entry)  
 DE Sequence of clone pVAM05R3 which contains probes for the in vitro  
 DE diagnosis of patients presenting an alteration on their chromosome  
 KW PCR primer: Charcot-Marie-Tooth type 1 disease; in vitro diagnosis;  
 KW chromosome 17p; ss.  
 OS Homo sapiens.







KW Barley: heat resistant; promoter: genetic engineering; transgenic;  
 PL Plant; improved maturation; ds.  
 PR WO9703133-A.  
 PD 23-JAN-1997.  
 PF 05-JUL-1996; J01866.  
 PI Ito K, Kihara M, Okada T, Yoshigi N;  
 DR WO/97/10896/20.  
 PT Expression promoter for genes inserted into plant seeds - e.g.  
 PT heat-resistant beta-amylase gene inserted into barley seeds.  
 CC T6352 is a DNA molecule comprising a heat resistant promoter derived  
 CC from a barley beta-amylase (EC 3.2.1.2) gene and other sequences  
 CC used for integration of this promoter into a reporter plasmid. The  
 CC plasmids are used for the production of transgenic barley plants  
 CC and seeds. 2142 BP; 551 A; 482 C; 395 G; 614 T.  
 SQ Sequence 2142 BP; 551 A; 482 C; 395 G; 614 T.

Query Match 1.34; Score 14; DB 33; Length 2142;  
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 327 attcttcctccca 340  
 Cp 994 ATTCTTCTCTCA 991

RESULT 329  
 ID T42218 standard; cDNA; 2152 BP.  
 AC T79594.  
 DT 08-OCT-1997 (first entry)  
 DE Human TATA-binding protein associated factor, hTAP1100 cDNA.  
 KW TATA-binding protein associated factor; TAP; nuclear protein;  
 KW TATA-binding protein associated factor; TAP; nuclear protein;  
 OS Homo sapiens.  
 PE Key Location/Qualifiers  
 FT cds 1..2115  
 FT /note= "no start codon"  
 FT US5637686-A.  
 PD 10-JUN-1997.  
 PI Comai L, Dynlacht BD, Heery T, Ruppert S, Tanese N;  
 DR WO/97/10896/20.  
 PT Expression promoter for genes inserted into plant seeds - e.g.  
 PT heat-resistant beta-amylase gene inserted into barley seeds.  
 CC T6352 is a DNA molecule comprising a heat resistant promoter derived  
 CC from a barley beta-amylase (EC 3.2.1.2) gene and other sequences  
 CC used for integration of this promoter into a reporter plasmid. The  
 CC plasmids are used for the production of transgenic barley plants  
 CC and seeds. 2142 BP; 551 A; 482 C; 395 G; 614 T.  
 SQ Sequence 2142 BP; 551 A; 482 C; 395 G; 614 T.

PI Tjian R, Wang E, Weinsierl ROJ;  
 DR WO/97/10896/20.  
 PT Nucleic acids encoding human TATA-binding protein associated factor  
 PT for modulating transcription of TAPs.  
 CC T79594 encodes human TATA-binding protein associated factor (TAP)  
 CC cDNA. The TAP cDNA is 2152 bp, 551 A, 482 C, 395 G, 614 T.  
 CC TAP is a nuclear protein involved in RNA polymerase II,  
 CC hTAP1100, hTAP1250, hTAP148 and hTAP1100 and nucleic acids encoding  
 CC them, are used to modulate transcription, including transcription  
 CC initiation. TAPs are nuclear proteins involved in RNA polymerase II,  
 CC hTAP1100, hTAP1250, hTAP148 and hTAP1100 and nucleic acids encoding  
 CC TAP, an activator, or TBP (TATA-binding protein) competitively  
 CC inhibiting association of a TAP domain with another compound, typically  
 CC a protein like TBP or another TAP, an activator, or DNA.  
 SQ Sequence 2152 BP; 551 A; 482 C; 395 G; 614 T.

Query Match 1.34; Score 15; DB 32; Length 2152;  
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1505 attattttgtctag 1519  
 Qy 22 ATTATTGTCTAG 36

RESULT 328  
 ID Q70732 standard; cDNA; 2152 BP.  
 AC Q70732.  
 DT 27-MAR-1995 (first entry)  
 DE TATA-binding protein associated factor, hTAP1100 cDNA.  
 KW TATA-binding protein associated factor; hTAP1100; as. screening;  
 KW diagnostic; therapeutic; gene transcription regulation.  
 OS Homo sapiens.  
 PE Key Location/Qualifiers  
 FT cds 1..2115  
 FT /note= "no start codon"  
 FT US5637686-A.  
 PD 10-JUN-1997.  
 PI Comai L, Dynlacht BD, Heery T, Ruppert S, Tanese N;  
 DR WO/97/10896/20.  
 PT Expression promoter for genes inserted into plant seeds - e.g.  
 PT heat-resistant beta-amylase gene inserted into barley seeds.  
 CC T6352 is a DNA molecule comprising a heat resistant promoter derived  
 CC from a barley beta-amylase (EC 3.2.1.2) gene and other sequences  
 CC used for integration of this promoter into a reporter plasmid. The  
 CC plasmids are used for the production of transgenic barley plants  
 CC and seeds. 2142 BP; 551 A; 482 C; 395 G; 614 T.  
 SQ Sequence 2142 BP; 551 A; 482 C; 395 G; 614 T.

PT in screening, diagnostics and therapeutics  
 PS Disclosure: Page 143-147; 180pp; English.  
 CC The TATA-binding protein associated factor hTAP1100 (including  
 CC cDNA, RNA polymerase II, the RNA polymerase which transcribes messenger  
 CC RNA, RNA polymerase II transcription process in vitro upon addition  
 CC of specific biochemical assays for screening compounds that agonize or  
 CC antagonize selected transcription factors involved in regulating  
 CC gene expression associated with human pathology.  
 SQ Sequence 2152 BP; 551 A; 482 C; 395 G; 614 T.  
 Query Match 1.34; Score 15; DB 12; Length 2152;  
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1505 attattttgtctag 1519  
 Qy 22 ATTATTGTCTAG 36

RESULT 329  
 ID T42218 standard; cDNA; 2152 BP.  
 AC T79594.  
 DT 08-OCT-1997 (first entry)  
 DE Human TATA-binding protein associated factor hTAP1100 gene.  
 KW Human TATA-binding protein; TBP associated factor; TAP; holoenzyme;  
 KW RNA polymerase II; transcription; messenger RNA; nuclear fraction;  
 OS Homo sapiens.  
 PE Key Location/Qualifiers  
 FT cds 1..2115  
 FT /note= "no start codon"  
 FT US5637686-A.  
 PD 10-JUN-1997.  
 PI Comai L, Dynlacht BD, Heery T, Ruppert S, Tanese N;  
 DR WO/97/10896/20.  
 PT Expression promoter for genes inserted into plant seeds - e.g.  
 PT heat-resistant beta-amylase gene inserted into barley seeds.  
 CC T6352 is a DNA molecule comprising a heat resistant promoter derived  
 CC from a barley beta-amylase (EC 3.2.1.2) gene and other sequences  
 CC used for integration of this promoter into a reporter plasmid. The  
 CC plasmids are used for the production of transgenic barley plants  
 CC and seeds. 2142 BP; 551 A; 482 C; 395 G; 614 T.  
 SQ Sequence 2142 BP; 551 A; 482 C; 395 G; 614 T.

CC polymerase II in vitro transcription activity. The encoded protein has  
 CC an estimated mol. wt. of 100 kD by SDS-PAGE.  
 CC The invention relates to purified proteins involved in transcription  
 CC by RNA polymerase II, the RNA polymerase which transcribes messenger  
 CC RNA, RNA polymerase II transcription process in vitro upon addition  
 CC of specific biochemical assays for screening compounds that agonize or  
 CC antagonize selected transcription factors involved in regulating  
 CC gene expression associated with human pathology.  
 SQ Sequence 2152 BP; 551 A; 482 C; 395 G; 614 T.  
 Query Match 1.34; Score 15; DB 25; Length 2152;  
 Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1505 attattttgtctag 1519  
 Qy 22 ATTATTGTCTAG 36

RESULT 330  
 ID T42218 standard; cDNA; 2152 BP.  
 AC T79594.  
 DT 22-JUN-1994 (first entry)  
 DE Toxoplasma GP28.5 antigen coding sequence.  
 KW Toxoplasma; Toxoplasma; tachyzoite; excystation; secretion antigen;  
 KW Toxoplasma; Toxoplasma; tachyzoite; excystation; secretion antigen;  
 OS Homo sapiens.  
 PE Key Location/Qualifiers  
 FT exon 986..1035  
 FT intron 1036..1274  
 FT exon 1275..1682  
 FT /tag= a  
 FT /tag= b  
 FT /tag= c  
 PD 23-DEC-1993.  
 PI Comai L, Dynlacht BD, Heery T, Ruppert S, Tanese N;  
 DR WO/97/10896/20.  
 PT Expression promoter for genes inserted into plant seeds - e.g.  
 PT heat-resistant beta-amylase gene inserted into barley seeds.  
 CC T6352 is a DNA molecule comprising a heat resistant promoter derived  
 CC from a barley beta-amylase (EC 3.2.1.2) gene and other sequences  
 CC used for integration of this promoter into a reporter plasmid. The  
 CC plasmids are used for the production of transgenic barley plants  
 CC and seeds. 2142 BP; 551 A; 482 C; 395 G; 614 T.  
 SQ Sequence 2142 BP; 551 A; 482 C; 395 G; 614 T.



Claim 21: Page 23-25; 37/99; French.  
 A T gondii genomic library in  $\lambda$ EMBL3 was screened with a fragment of  
 a clone isolated by immunoscreening a T gondii strain RH cDNA  
 library with a rabbit anti-T gondii tachyzoite antigen. The clone  
 encodes the complete 38 kDa antigen was identified. The antigen is  
 the principal excretion-secretion antigen of T gondii tachyzoites  
 and represents a good candidate for effective vaccine production.  
 Sequence 235 Aa; 335 Aa; 355 C; 397 T.  
 Query Match: Score 14; DB 9; Length 2152;  
 Local Similarity 100.0%; Percent 14.78+02;  
 Matrix: 4, Conservative; 0, Mismatches; 0, Indels; 0, Gaps; 0,

Accession	Result	331	Standard	DNA	2154 BP.
U0125	AC	664125	AC	664125	AC
U0126	AC	664126	AC	664126	AC
U0127	AC	664127	AC	664127	AC
U0128	AC	664128	AC	664128	AC
U0129	AC	664129	AC	664129	AC
U0130	AC	664130	AC	664130	AC
U0131	AC	664131	AC	664131	AC
U0132	AC	664132	AC	664132	AC
U0133	AC	664133	AC	664133	AC
U0134	AC	664134	AC	664134	AC
U0135	AC	664135	AC	664135	AC
U0136	AC	664136	AC	664136	AC
U0137	AC	664137	AC	664137	AC
U0138	AC	664138	AC	664138	AC
U0139	AC	664139	AC	664139	AC
U0140	AC	664140	AC	664140	AC
U0141	AC	664141	AC	664141	AC
U0142	AC	664142	AC	664142	AC
U0143	AC	664143	AC	664143	AC
U0144	AC	664144	AC	664144	AC
U0145	AC	664145	AC	664145	AC
U0146	AC	664146	AC	664146	AC
U0147	AC	664147	AC	664147	AC
U0148	AC	664148	AC	664148	AC
U0149	AC	664149	AC	664149	AC
U0150	AC	664150	AC	664150	AC
U0151	AC	664151	AC	664151	AC
U0152	AC	664152	AC	664152	AC
U0153	AC	664153	AC	664153	AC
U0154	AC	664154	AC	664154	AC
U0155	AC	664155	AC	664155	AC
U0156	AC	664156	AC	664156	AC
U0157	AC	664157	AC	664157	AC
U0158	AC	664158	AC	664158	AC
U0159	AC	664159	AC	664159	AC
U0160	AC	664160	AC	664160	AC
U0161	AC	664161	AC	664161	AC
U0162	AC	664162	AC	664162	AC
U0163	AC	664163	AC	664163	AC
U0164	AC	664164	AC	664164	AC
U0165	AC	664165	AC	664165	AC
U0166	AC	664166	AC	664166	AC
U0167	AC	664167	AC	664167	AC
U0168	AC	664168	AC	664168	AC
U0169	AC	664169	AC	664169	AC
U0170	AC	664170	AC	664170	AC
U0171	AC	664171	AC	664171	AC
U0172	AC	664172	AC	664172	AC
U0173	AC	664173	AC	664173	AC
U0174	AC	664174	AC	664174	AC
U0175	AC	664175	AC	664175	AC
U0176	AC	664176	AC	664176	AC
U0177	AC	664177	AC	664177	AC
U0178	AC	664178	AC	664178	AC
U0179	AC	664179	AC	664179	AC
U0180	AC	664180	AC	664180	AC
U0181	AC	664181	AC	664181	AC
U0182	AC	664182	AC	664182	AC
U0183	AC	664183	AC	664183	AC
U0184	AC	664184	AC	664184	AC
U0185	AC	664185	AC	664185	AC
U0186	AC	664186	AC	664186	AC
U0187	AC	664187	AC	664187	AC
U0188	AC	664188	AC	664188	AC

09-11-1991 A.  
08-OCT-1993: U09636.  
25-NOV-1992: US-980518.  
1989M) BRIGHAN & WOMENS HOSPITAL.  
Schubach W, Kieff E.  
1989; 24.  
P-PSDB: 554079.  
DNA coding for Epstein Barr virus induced (EBI) polypeptide(s)  
and antibodies to EMI. 2 and 3 - useful for detecting EBV by  
hybridisation or by immunoassay  
EBV in EBV infected lymphocytes stimulate features of antigen  
activation in enlarging, increasing RNA synthesis, expressing  
activation antigens and adhesion molecules, secreting Ig and  
proliferating. Unlike antigen stimulated B lymphocytes, EBV  
infected B lymphocytes continue to proliferate (in vitro) after  
effects of EBV and antigen induced genes are likely to include  
mediators of antigen induced B lymphocyte growth or differentiation.

Tue Nov 17 08:55:26 1998

**Tue Nov 17 08:55:26 1998**

Page 222

[illegible]

KEY	CDS	Location/Qualifiers
CC	NC_009746.1	/stage = 8
CC	N09746684-A1..	
CC	11-DEC-1997.	
CC	U09011.	
CC	06-JUN-1997.	
CC	11-DEC-1997.	
CC	U09011.	
CC	(UTM) UNIV MASSACHUSETTS,	
CC	Davis RJ, Galicheva-Gargova Z;	
CC	WPI : 98-04132/04.	
CC	P-PSDB : N29494.	
CC	Abstract:	
CC	transcript encoding RNA-binding polypeptide zfp1 - useful	
CC	to reduce angiogenesis, increase cell proliferation or kidney cell	
CC	regeneration or inhibit tumour growth	
CC	Claim 4: Fig 1: 8pp; English.	
CC	This nucleotide sequence comprises a cDNA clone coding for mouse	
CC	zfp1, a novel class of RNA binding proteins. The blood tissue	
CC	signalling molecule, a novel protein kinase C (PKC)-beta	
CC	-activated receptors (e.g. EGF and PDGF receptors)	
CC	specifically bind small nuclear RNAs (e.g. U3).	
CC	A yeast two-hybrid system was used to identify proteins that bind to the	
CC	cysteine of non-activated PKC beta1b and several fragments	
CC	were identified. A murine foetal cell library was screened and	
CC	used to probe a murine foetal cell library to isolate the full	
CC	length zfp1 sequence. This was used as a probe to isolate	
CC	homologous sequences (see V04062-4) from Saccharomyces cerevisiae,	
CC	Saccharomycoccus pombe and human. The isolated clones can	
CC	complement the growth defect of the yeast strains expressing	
CC	elevated level in tumour derived cells, and is a marker for growth	
CC	and differentiation processes, including malignant transformation	
CC	of cells. zfp1, optionally expressed by gene therapy methods, can	
CC	be used to inhibit the proliferation of zfp1 associated malignant	
CC	tumours. zfp1 expression suppresses increase cell proliferation,	
CC	reduce angiogenesis, e.g. to treat tumours, diabetic retinopathy,	

```

5Q Sequence 2154 BP; 486 A; 647 C; 543 G; 478 T;
Query Match 1.38; Score 14; DB 11; Length 2154;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 370 gscagacatcctctt 383
|||||
Oy 244 SCAGACATCCTCTT 257

BSRSTG 323
ID SC6160 standard; cDNA; 2160 BP.
AC G66160;
O 0-P28-1995 (first entry)
DE Putative seven transmembrane receptor (V13-B) coding sequence (cDNA).
PR Primer: seven transmembrane receptor; receptor: amplification; PCR;
KW Homo sapiens.
OS Homo sapiens.

```

PR	Key	Location/Qualifiers
PR	Cds	74620
PR		/+acc
PR		/product= Seven transmembrane receptor V31.
PR		W09412635: A.
PR		19-JUN-1994.
PR		17-NOV-1992; U11153
PR		17-NOV-1992; US-977452.
PR		(ICGS) ICOS CORP.
PA		Godiska R, Gray PW, Schweikart VL,
PA		Godiska R, Gray PW, Schweikart VL.
PR		P-DB08: 853774.
PR		DR
PT		DNA encoding seven trans:membrane receptors - used to develop
PT		prods. for use as therapeutic or diagnostic agents
PT		involving 3 Peptides: 100pp; English.
PT		Abstract 3. Peptides: 100pp; English.
CC		A human cDNA encoding the seven transmembrane receptor V31 was
CC		isolated by first amplifying a partial cDNA clone from a human
CC		testis cDNA library using two primers (066150, 066155). The resulting
CC		sequence (066157) showed a hybridizing band was isolated from the
CC		sequences (066155, 066157) probes. A partial cDNA clone was
CC		941 and cloned. The resulting clone was named pV31.5'end (066158). A
CC		full length cDNA clone was isolated from a peripheral blood
CC		mononuclear cell library using V31 specific primers (066159, 066152).
CC		CDC (PNCMS) V31-5'end using V31 cDNA insert in the clone was
CC		CC designated CNA V31-5' (066160).
CC		Sequence 2160 BP: 645 C: 546 G: 479 T:
CC		Query Match 133; Score 14; DP 11; Length 2160;
CC		Best Local Similarity 100.00; Pred. No. 3,477+02; Indels 0;
CC		Matches 14; Conservative 0; Mismatches 0; Gaps 0;

CC Pharyngitis, arthritis, proctitis and coronary atherosclerosis  
CC increase tubular regeneration of kidney cells, e.g.: following acute  
CC renal failure.  
CC Sequence 2162 BP: 607 A: 573 C: 571 G: 411 T:  
Query Match 1.3% Score 14: DB 39: Length 2162:  
Best Local Similarity 100.0% Pred. No: 34+02:  
Matches 14: Conservative 0: Mismatches 0: Indels 0: Gaps 0:  
DB 1642 taccgtgctctc 1655

```

73 DCTGTCGTCCCTC 86

RESULT 334
ID T14054 standard; cDNA; 2165 BP.
DT 1995-01-11
DV 20-MOV-1996 (first entry)
DE Fungal signal recognition particle, SRP54 gene cDNA clone.
KW Signal recognition particle; recombinant; genetic engineering;
NM functional domains; fungi; transformation; Aspergillus;
NS Aspergillus niger.
FE Key Location/Qualifiers
FT CD      86..1693 /cdseq = a
PR NO5610635-A.
FD 11-APR-1996.
PP PP      02-OCT-1995; UI2628.
PA OS OR    OJ00000001
PI NC       NOVATO BIOTECH INC.
PL Thompson SA, Zaver DS.
PT WP1: 96-209357/21.
PW P-WPB: R93130.
DR     Amino acid construct encoding elements of signal recognition particle
DS     from various fungi - used to increase prodn. of heterologous proteins
in fungi.
PC Claim 8, fig 3. Jpp: English.
CC T14053 and T14054 both encode an Aspergillus niger signal recognition
particle. The amino acids encoded by each cDNA are identical except at CC
SPK23 clone respectively. A nucleic acid construct containing DNA coding
for the entire protein sequence as shown above will produce a protein of
interest, in particular for the recombinant production of a protein of
commercial interest, such as xylanase, cellulase, glucoamylase, pectinase or
other enzymes useful in food processing. This may include other components
of transformed cells, improving the production of secreted
proteins.
CO Proteins.
EQ Query Match          1.3%   Score 14; DB 23; Length 2165.
Best Local similarity 100.0%; Pred. No. 3,47+02.
```



	Matches	14: Conservative	0: Mismatches	0: Indels	0: Caps
Db	665	cgtcaccaagcaggga	678		
Cp	804	CCTCACAAGCAGGA	791		

RESUME 335  
 ID: N82402 standard; DMA. 2177 BP.  
 AC N82402:  
 DT 26-Nov-1990 (first entry)  
 DE b subunit of human Factor XIII.  
 DS b subunit of factor XIII; blood clot stabilization; fibrin polymer;  
 OS crosslinking; fibrinogen; fibrinogen; fibrinogen; fibrinogen;  
 OS Homo sapiens. Location/Qualifiers  
 Key words  
 cds 2..1521  
 tagseq\_a 2..1521  
 tagseq\_b subunit 2..58  
 misc\_rna 2..58  
 /tagseq\_b  
 /label=leader\_sequence  
 58..1521  
 /product=mature b subunit  
 /cdd=1  
 AU8778694-A.  
 31-MAR-1988. 078694  
 13-SEP-1986. U2-009512  
 (F8MO.) ZymoGenetics Inc.  
 PA Davis E, Seale RL, Ichinose A, Holly JM, Parker GE.;  
 WP1: 89-140637/21.  
 NEW ENGL J MED 321:1221-1225 (1989) 1221-1225  
 DT 19890621  
 PR expression vectors encoding sub-units of factor 13 - and corresponding  
 PR expression vectors and transformed host cells.  
 PR Disclosure: P: English.  
 CC The carboxyl-terminal Thr (nucleotides 1979-1981) is followed by  
 CC a 19 bp noncoding sequence, and a poly(A) tail  
 CC of 5 bp codon (TAG), a 19 bp noncoding sequence, and a poly(A) tail  
 CC of 5 bp codon (TAG).  
 CC was identified 19 nucleotides upstream from the poly(A) tail.  
 CC When cultured the host cells will produce the polypeptide which  
 CC can be assembled to factor 13, which stabilizes blood clots by  
 CC crosslinking fibrin polymers. Factor 13 is useful therapeutically, e.g.  
 CC in the treatment of hemophilia.  
 CC can be prep'd in large amounts without risk of viral contamination.  
 CC See also N82401 and N82403.  
 CC Sequence 2177 BP: 758 A: 358 C: 434 G: 627 T:  
 133: Score 14: DB 1: Length 2177;  
 Best Local similarity 100.0%: Pred. No. 3,474+02;  
 Matches 14: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Tue Nov 17 08:55:26 1998

US-08-887-977-9.rmg

Best Local Similarity 100.0%;  
Matches 14: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 327
ID ID Q10958 standard: DNA: 2196 BP.
AC Q10958
UT 21-MAY-1991 (first entry)
DE Gene encoding acylamino acid-isolating enzyme-like polypeptide.
RW ARE; pig liver; ss.
RT Genes scrofa.
ES Location/Qualifiers
FT cds
FT 1..2196
FT /tag= a
FT /product= AARE-like polypeptide
FT 10-030673.A.
FT 29-JUN-1989.
FT 29-JUN-1989: 165216.
FT 29-JUN-1989: JP-155216.
FT (TAKA-) TAKARA SHIZO KK.
FT WPI: 11-084340/12.
FT P-9DB; R1008-isolating enzyme-like polypeptide - prepd. by
FT genetic engineering; Japanese.
FT Claim 1: Fig 2: 10pp: Japanese.
FT The sequence was obt'd. from six clones isolated from a pig liver
FT cDNA library. The clones, lambda AARE 459, 450, 451, 452, 521 and
FT 122 cover a large sequence of the gene.
FT Sequence 2196 BP: 438 A. 594 C. 677 G. 487 T.
Query Match
Best Local Similarity 100.0% Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Gaps 0; Gaps 0;
ob 1216 tcagctggaaagctg 1229
|||||
27 596 tcagctggaaagctg 609

RESULT 338
ID ID Q49032 standard: DNA: 2198 BP.
AC Q49032
UT 22-APR-1994 (first entry)
DE Granulin coding sequence.
RW 2E granulin; keratinocytes; wound healing; inhibition; peptide;
RT Genes scrofa; granulocytes; leukocytes; ss.
ES Homo sapiens.

```

**Tue Nov 17 08:55:26 1998**

Key	Location/Qualifiers
misc_difference	1453..1463
/tag=	a
/transl_except=	ATG
/tag=	16
misc_difference	/tag=
/transl_except=	CAG
MO313195.A1	
28-FEB-1992;	CA0089.
03-FEB-1992;	DS-829233.
(SOLFO)/	SOLFOH.S
WPI:	33.320328/40.
P-PDB:	R48673.
New cysteine-rich granulin peptide	
Disclature: Figure 4; 53pp; English	
The granulin inhibits keratinocytes	
in the epidermis of the mouse	
Sequence	238 Bp, 70%
Query Match	Similarity 10.1%
Matches	14; Conservative
Db	928 cyscatcactgctgc 841
QY	991 ccsacacacacacac 904

[illegible]

Page 336

```

      s Valine.
      s glycine.
      leucocyte(s) - are
      useful for wound healing
      useful in formulations
      646 G; 446 T;
      8; Length 2198;
      s 0; Index 0; Caps 0;

```

ase coding sequence.  
photosynthesis; cordgrass;  
ature; probe; Panicum maximum;  
ss.  
rate carboxykinase











```

RESULT 349
ID T16816 standard; DNA; 2259 BP.
AC T16816; 1996 (first entry)
DE Kaposi's sarcoma associated herpesvirus ORP24.
KW Kaposi's sarcoma; KSHV; lymphoma; AIDS; vaccine; diagnosis; therapy;
KW es.
KW Kaposi's sarcoma associated herpesvirus.
AC W98F060159-A1.
PD 29-FEB-1996.
PF 11-AUG-1985; D10194.
PF 21-AUG-1984; US-202165.
PF 11-AUG-1984; US-202165.
PR 11-APR-1995; US-420315.
PR (UTIC) UNIV COLOMBIA NEW YORK.
PR (GRAMV) GRANT D E.
PA (HIV) HANSEN T.
PI Chang Y, Moore PS, Grant DE, Viole L.
PI WPI: 96-151362/15.
PR P-PSDB: R03610.
DR Herpesvirus DNA associated with Kaposi's sarcoma - also associated
DR with lymphoma and vaccination.
PS Claim 17: Page 220-222; 30Sep; English.
CC Kaposi's sarcoma associated herpes virus (KSHV) clone K85 (T16806), an
CC optd. from a KS lesion genomic library, includes 15 complete ORFs and
CC encodes a protein homologous to the gamma 1 and gamma 2 proteins of the
CC salivary positional homologues. ORP24 is given in T16816 and the
CC encoded protein in R03610. KSHV DNA may be incorporated into a
CC vector and expressed in host cells to produce peptides useful in
CC immunisation or as a vaccine. The protein itself may be used to
CC immunise subjects or used to design the specific antigenic
CC triplex molecules or diagnostic probes, or to raise transgenic
CC animals.
SQ Sequence 2259 BP; 586 A; 632 C; 551 G; 490 T;
Query Match 1.38; Score 15; DB 20; Length 2259;
Best Local Similarity 100.0%; Prid No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps
0
Db 676 sacatcncagcagcttc 690
|||||
Cc 617 MACATCAGCAGCTTC 603

RESULT 350
ID T30686 standard; DNA; 2259 BP.
AC T30686; 1996 (first entry)
DE Kaposi's sarcoma associated herpesvirus ORP24.
KW Kaposi's sarcoma; KSHV; lymphoma; AIDS; vaccine; diagnosis; therapy;
KW es.
KW Kaposi's sarcoma associated herpesvirus.
AC W98F060159-A1.
PD 29-FEB-1996.
PF 11-AUG-1985; D10194.
PF 21-AUG-1984; US-202165.
PF 11-AUG-1984; US-202165.
PR 11-APR-1995; US-420315.
PR (UTIC) UNIV COLOMBIA NEW YORK.
PR (GRAMV) GRANT D E.
PA (HIV) HANSEN T.
PI Chang Y, Moore PS, Grant DE, Viole L.
PI WPI: 96-151362/15.
PR P-PSDB: R03610.
DR Herpesvirus DNA associated with Kaposi's sarcoma - also associated
DR with lymphoma and vaccination.
PS Claim 17: Page 220-222; 30Sep; English.
CC Kaposi's sarcoma associated herpes virus (KSHV) clone K85 (T16806), an
CC optd. from a KS lesion genomic library, includes 15 complete ORFs and
CC encodes a protein homologous to the gamma 1 and gamma 2 proteins of the
CC salivary positional homologues. ORP24 is given in T16816 and the
CC encoded protein in R03610. KSHV DNA may be incorporated into a
CC vector and expressed in host cells to produce peptides useful in
CC immunisation or as a vaccine. The protein itself may be used to
CC immunise subjects or used to design the specific antigenic
CC triplex molecules or diagnostic probes, or to raise transgenic
CC animals.
SQ Sequence 2259 BP; 586 A; 632 C; 551 G; 490 T;
Query Match 1.38; Score 15; DB 20; Length 2259;
Best Local Similarity 100.0%; Prid No. 8.37e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps
0
Db 676 sacatcncagcagcttc 690
|||||
Cc 617 MACATCAGCAGCTTC 603

```

```

PF 10-NOV-1995; DK048.
PA (NOVO) NOVO-NORDISK AS.
PI Bjorbaek C, Enevold SM, Pedersen O;
PR wpi; 96-259840/26.
DR Mutant DNA encoding insulin-stimulated protein kinase - useful to
PS diagnose pre-disposition to insulin resistance and possibly MIDDH
PT Claim 3; Page 30; 42pp; English.
PC DNA encoding a mutant insulin-stimulated protein kinase (ISPK-1)
CC genes. The method of detection comprises analysing an isolated DNA
CC by digestion with a restriction endonuclease which cleaves at the
CC site of the mutation. The obtained restriction pattern is compared
CC with a negative control restriction pattern (obtained from a wild type
CC ISPK-1 gene) to determine the presence of the mutation. The method
CC from a mutant ISPK-1 gene. This method may be used to diagnose
CC pre-disposition to insulin resistance and possibly non-insulin-
CC dependent diabetes mellitus. This mutant sequence contains a T to
CC C substitution in the second codon 38 resulting in a
CC sequence 2260 BP: 704 A: 417 C: 511 G: 628 T;
SQ
Query Match 138; Score 14; DB 23; Length 2260;
Best Local Similarity 100.0%; Prod No. 347e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps
Db 1346 carytagattatt 1359
QY 14 CACTGAAGATTATT 27

RESULT 352
ID T27730 standard; DNA; 2260 BP.
ID T27730;
DT 08-NOV-1996 (first entry)
DT 10-NOV-1995
DT Insulin-stimulated protein kinase (ISPK-1) coding sequence
KW Insulin-stimulated protein kinase; diabetes mellitus; screening;
KW diagnosis; resistance; diabetes mellitus; MIDDH; ss.
KW Homo sapiens.
FH Key Location/Qualifiers
FH Key 1359
FH CDS 1353
FH /translation
FH /product= Insulin-stimulated protein kinase.
PF W09615233.A1.
PA 10-NOV-1995
PI 10-NOV-1995; DK048.
PR 10-NOV-1994; DK-001294.
PA (NOVO) NOVO-NORDISK AS.
PI Bjorbaek C, Enevold SM, Pedersen O;

```











CC Same SacB[*lac*] probes were used in the Southern hybridization and  
 CC this hybridized only with a 1.5kb fragment. Sequence analysis of  
 CC this fragment revealed that the gene was identical to the mouse  
 CC regulatory region can be incorporated into stable vectors to  
 CC control expression of heterologous genes in transformed *E. coli*.  
 CC See also Q10497-7.

CC Query Match 1.38; Score 14; DB 2; Length 2351;  
 CC Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC Db 2208 atgacatcagag 2221  
 CC Cp 662 ATGACATCAGAG 649

CC RESULT 361

CC ID T97972 standard; DNA: 2379 BP.

CC AC T97972 (first entry)

CC DE Nucleotide sequence of DEN-1 16007 PDK-13 vaccine virus.

CC KW DEN-1 virus; polyprotein: capsid; PRM; M; E; NS1; NS2A; NS2B; NS3;  
 CC NS4; NS5; NS6; NS7; NS8; NS9; NS10; NS11; NS12; NS13; NS14;  
 CC NS15; NS16; NS17; NS18; NS19; NS20; NS21; NS22; NS23; NS24;  
 CC NS25; NS26; NS27; NS28; NS29; NS30; NS31; NS32; NS33; NS34;  
 CC NS35; NS36; NS37; NS38; NS39; NS40; NS41; NS42; NS43; NS44;  
 CC NS45; NS46; NS47; NS48; NS49; NS50; NS51; NS52; NS53; NS54;  
 CC NS55; NS56; NS57; NS58; NS59; NS60; NS61; NS62; NS63; NS64;  
 CC NS65; NS66; NS67; NS68; NS69; NS70; NS71; NS72; NS73; NS74;  
 CC NS75; NS76; NS77; NS78; NS79; NS80; NS81; NS82; NS83; NS84;  
 CC NS85; NS86; NS87; NS88; NS89; NS90; NS91; NS92; NS93; NS94;  
 CC NS95; NS96; NS97; NS98; NS99; NS100; NS101; NS102; NS103;  
 CC NS104; NS105; NS106; NS107; NS108; NS109; NS110; NS111;  
 CC NS112; NS113; NS114; NS115; NS116; NS117; NS118; NS119;  
 CC NS120; NS121; NS122; NS123; NS124; NS125; NS126; NS127;  
 CC NS128; NS129; NS130; NS131; NS132; NS133; NS134; NS135;  
 CC NS136; NS137; NS138; NS139; NS140; NS141; NS142; NS143;  
 CC NS144; NS145; NS146; NS147; NS148; NS149; NS150; NS151;  
 CC NS152; NS153; NS154; NS155; NS156; NS157; NS158; NS159;  
 CC NS160; NS161; NS162; NS163; NS164; NS165; NS166; NS167;  
 CC NS168; NS169; NS170; NS171; NS172; NS173; NS174; NS175;  
 CC NS176; NS177; NS178; NS179; NS180; NS181; NS182; NS183;  
 CC NS184; NS185; NS186; NS187; NS188; NS189; NS190; NS191;  
 CC NS192; NS193; NS194; NS195; NS196; NS197; NS198; NS199;  
 CC NS200; NS201; NS202; NS203; NS204; NS205; NS206; NS207;  
 CC NS208; NS209; NS210; NS211; NS212; NS213; NS214; NS215;  
 CC NS216; NS217; NS218; NS219; NS220; NS221; NS222; NS223;  
 CC NS224; NS225; NS226; NS227; NS228; NS229; NS230; NS231;  
 CC NS232; NS233; NS234; NS235; NS236; NS237; NS238; NS239;  
 CC NS240; NS241; NS242; NS243; NS244; NS245; NS246; NS247;  
 CC NS248; NS249; NS250; NS251; NS252; NS253; NS254; NS255;  
 CC NS256; NS257; NS258; NS259; NS260; NS261; NS262; NS263;  
 CC NS264; NS265; NS266; NS267; NS268; NS269; NS270; NS271;  
 CC NS272; NS273; NS274; NS275; NS276; NS277; NS278; NS279;  
 CC NS280; NS281; NS282; NS283; NS284; NS285; NS286; NS287;  
 CC NS288; NS289; NS290; NS291; NS292; NS293; NS294; NS295;  
 CC NS296; NS297; NS298; NS299; NS300; NS301; NS302; NS303;  
 CC NS304; NS305; NS306; NS307; NS308; NS309; NS310; NS311;  
 CC NS312; NS313; NS314; NS315; NS316; NS317; NS318; NS319;  
 CC NS320; NS321; NS322; NS323; NS324; NS325; NS326; NS327;  
 CC NS328; NS329; NS330; NS331; NS332; NS333; NS334; NS335;  
 CC NS336; NS337; NS338; NS339; NS340; NS341; NS342; NS343;  
 CC NS344; NS345; NS346; NS347; NS348; NS349; NS350; NS351;  
 CC NS352; NS353; NS354; NS355; NS356; NS357; NS358; NS359;  
 CC NS360; NS361; NS362; NS363; NS364; NS365; NS366; NS367;  
 CC NS368; NS369; NS370; NS371; NS372; NS373; NS374; NS375;  
 CC NS376; NS377; NS378; NS379; NS380; NS381; NS382; NS383;  
 CC NS384; NS385; NS386; NS387; NS388; NS389; NS390; NS391;  
 CC NS392; NS393; NS394; NS395; NS396; NS397; NS398; NS399;  
 CC NS400; NS401; NS402; NS403; NS404; NS405; NS406; NS407;  
 CC NS408; NS409; NS410; NS411; NS412; NS413; NS414; NS415;  
 CC NS416; NS417; NS418; NS419; NS420; NS421; NS422; NS423;  
 CC NS424; NS425; NS426; NS427; NS428; NS429; NS430; NS431;  
 CC NS432; NS433; NS434; NS435; NS436; NS437; NS438; NS439;  
 CC NS440; NS441; NS442; NS443; NS444; NS445; NS446; NS447;  
 CC NS448; NS449; NS450; NS451; NS452; NS453; NS454; NS455;  
 CC NS456; NS457; NS458; NS459; NS460; NS461; NS462; NS463;  
 CC NS464; NS465; NS466; NS467; NS468; NS469; NS470; NS471;  
 CC NS472; NS473; NS474; NS475; NS476; NS477; NS478; NS479;  
 CC NS480; NS481; NS482; NS483; NS484; NS485; NS486; NS487;  
 CC NS488; NS489; NS490; NS491; NS492; NS493; NS494; NS495;  
 CC NS496; NS497; NS498; NS499; NS500; NS501; NS502; NS503;  
 CC NS504; NS505; NS506; NS507; NS508; NS509; NS510; NS511;  
 CC NS512; NS513; NS514; NS515; NS516; NS517; NS518; NS519;  
 CC NS520; NS521; NS522; NS523; NS524; NS525; NS526; NS527;  
 CC NS528; NS529; NS530; NS531; NS532; NS533; NS534; NS535;  
 CC NS536; NS537; NS538; NS539; NS540; NS541; NS542; NS543;  
 CC NS544; NS545; NS546; NS547; NS548; NS549; NS550; NS551;  
 CC NS552; NS553; NS554; NS555; NS556; NS557; NS558; NS559;  
 CC NS560; NS561; NS562; NS563; NS564; NS565; NS566; NS567;  
 CC NS568; NS569; NS570; NS571; NS572; NS573; NS574; NS575;  
 CC NS576; NS577; NS578; NS579; NS580; NS581; NS582; NS583;  
 CC NS584; NS585; NS586; NS587; NS588; NS589; NS590; NS591;  
 CC NS592; NS593; NS594; NS595; NS596; NS597; NS598; NS599;  
 CC NS600; NS601; NS602; NS603; NS604; NS605; NS606; NS607;  
 CC NS608; NS609; NS610; NS611; NS612; NS613; NS614; NS615;  
 CC NS616; NS617; NS618; NS619; NS620; NS621; NS622; NS623;  
 CC NS624; NS625; NS626; NS627; NS628; NS629; NS630; NS631;  
 CC NS632; NS633; NS634; NS635; NS636; NS637; NS638; NS639;  
 CC NS640; NS641; NS642; NS643; NS644; NS645; NS646; NS647;  
 CC NS648; NS649; NS650; NS651; NS652; NS653; NS654; NS655;  
 CC NS656; NS657; NS658; NS659; NS660; NS661; NS662; NS663;  
 CC NS664; NS665; NS666; NS667; NS668; NS669; NS670; NS671;  
 CC NS672; NS673; NS674; NS675; NS676; NS677; NS678; NS679;  
 CC NS680; NS681; NS682; NS683; NS684; NS685; NS686; NS687;  
 CC NS688; NS689; NS690; NS691; NS692; NS693; NS694; NS695;  
 CC NS696; NS697; NS698; NS699; NS700; NS701; NS702; NS703;  
 CC NS704; NS705; NS706; NS707; NS708; NS709; NS710; NS711;  
 CC NS712; NS713; NS714; NS715; NS716; NS717; NS718; NS719;  
 CC NS720; NS721; NS722; NS723; NS724; NS725; NS726; NS727;  
 CC NS728; NS729; NS730; NS731; NS732; NS733; NS734; NS735;  
 CC NS736; NS737; NS738; NS739; NS740; NS741; NS742; NS743;  
 CC NS744; NS745; NS746; NS747; NS748; NS749; NS750; NS751;  
 CC NS752; NS753; NS754; NS755; NS756; NS757; NS758; NS759;  
 CC NS760; NS761; NS762; NS763; NS764; NS765; NS766; NS767;  
 CC NS768; NS769; NS770; NS771; NS772; NS773; NS774; NS775;  
 CC NS776; NS777; NS778; NS779; NS780; NS781; NS782; NS783;  
 CC NS784; NS785; NS786; NS787; NS788; NS789; NS790; NS791;  
 CC NS792; NS793; NS794; NS795; NS796; NS797; NS798; NS799;  
 CC NS800; NS801; NS802; NS803; NS804; NS805; NS806; NS807;  
 CC NS808; NS809; NS810; NS811; NS812; NS813; NS814; NS815;  
 CC NS816; NS817; NS818; NS819; NS820; NS821; NS822; NS823;  
 CC NS824; NS825; NS826; NS827; NS828; NS829; NS830; NS831;  
 CC NS832; NS833; NS834; NS835; NS836; NS837; NS838; NS839;  
 CC NS840; NS841; NS842; NS843; NS844; NS845; NS846; NS847;  
 CC NS848; NS849; NS850; NS851; NS852; NS853; NS854; NS855;  
 CC NS856; NS857; NS858; NS859; NS860; NS861; NS862; NS863;  
 CC NS864; NS865; NS866; NS867; NS868; NS869; NS870; NS871;  
 CC NS872; NS873; NS874; NS875; NS876; NS877; NS878; NS879;  
 CC NS880; NS881; NS882; NS883; NS884; NS885; NS886; NS887;  
 CC NS888; NS889; NS890; NS891; NS892; NS893; NS894; NS895;  
 CC NS896; NS897; NS898; NS899; NS900; NS901; NS902; NS903;  
 CC NS904; NS905; NS906; NS907; NS908; NS909; NS910; NS911;  
 CC NS912; NS913; NS914; NS915; NS916; NS917; NS918; NS919;  
 CC NS920; NS921; NS922; NS923; NS924; NS925; NS926; NS927;  
 CC NS928; NS929; NS930; NS931; NS932; NS933; NS934; NS935;  
 CC NS936; NS937; NS938; NS939; NS940; NS941; NS942; NS943;  
 CC NS944; NS945; NS946; NS947; NS948; NS949; NS950; NS951;  
 CC NS952; NS953; NS954; NS955; NS956; NS957; NS958; NS959;  
 CC NS960; NS961; NS962; NS963; NS964; NS965; NS966; NS967;  
 CC NS968; NS969; NS970; NS971; NS972; NS973; NS974; NS975;  
 CC NS976; NS977; NS978; NS979; NS980; NS981; NS982; NS983;  
 CC NS984; NS985; NS986; NS987; NS988; NS989; NS990; NS991;  
 CC NS992; NS993; NS994; NS995; NS996; NS997; NS998; NS999;  
 CC NS1000; NS1001; NS1002; NS1003; NS1004; NS1005; NS1006;  
 CC NS1007; NS1008; NS1009; NS1010; NS1011; NS1012; NS1013;  
 CC NS1014; NS1015; NS1016; NS1017; NS1018; NS1019; NS1020;  
 CC NS1021; NS1022; NS1023; NS1024; NS1025; NS1026; NS1027;  
 CC NS1028; NS1029; NS1030; NS1031; NS1032; NS1033; NS1034;  
 CC NS1035; NS1036; NS1037; NS1038; NS1039; NS1040; NS1041;  
 CC NS1042; NS1043; NS1044; NS1045; NS1046; NS1047; NS1048;  
 CC NS1049; NS1050; NS1051; NS1052; NS1053; NS1054; NS1055;  
 CC NS1056; NS1057; NS1058; NS1059; NS1060; NS1061; NS1062;  
 CC NS1063; NS1064; NS1065; NS1066; NS1067; NS1068; NS1069;  
 CC NS1070; NS1071; NS1072; NS1073; NS1074; NS1075; NS1076;  
 CC NS1077; NS1078; NS1079; NS1080; NS1081; NS1082; NS1083;  
 CC NS1084; NS1085; NS1086; NS1087; NS1088; NS1089; NS1090;  
 CC NS1091; NS1092; NS1093; NS1094; NS1095; NS1096; NS1097;  
 CC NS1098; NS1099; NS1100; NS1101; NS1102; NS1103; NS1104;  
 CC NS1105; NS1106; NS1107; NS1108; NS1109; NS1110; NS1111;  
 CC NS1112; NS1113; NS1114; NS1115; NS1116; NS1117; NS1118;  
 CC NS1119; NS1120; NS1121; NS1122; NS1123; NS1124; NS1125;  
 CC NS1126; NS1127; NS1128; NS1129; NS1130; NS1131; NS1132;  
 CC NS1133; NS1134; NS1135; NS1136; NS1137; NS1138; NS1139;  
 CC NS1140; NS1141; NS1142; NS1143; NS1144; NS1145; NS1146;  
 CC NS1147; NS1148; NS1149; NS1150; NS1151; NS1152; NS1153;  
 CC NS1154; NS1155; NS1156; NS1157; NS1158; NS1159; NS1160;  
 CC NS1161; NS1162; NS1163; NS1164; NS1165; NS1166; NS1167;  
 CC NS1168; NS1169; NS1170; NS1171; NS1172; NS1173; NS1174;  
 CC NS1175; NS1176; NS1177; NS1178; NS1179; NS1180; NS1181;  
 CC NS1182; NS1183; NS1184; NS1185; NS1186; NS1187; NS1188;  
 CC NS1189; NS1190; NS1191; NS1192; NS1193; NS1194; NS1195;  
 CC NS1196; NS1197; NS1198; NS1199; NS1200; NS1201; NS1202;  
 CC NS1203; NS1204; NS1205; NS1206; NS1207; NS1208; NS1209;  
 CC NS1210; NS1211; NS1212; NS1213; NS1214; NS1215; NS1216;  
 CC NS1217; NS1218; NS1219; NS1220; NS1221; NS1222; NS1223;  
 CC NS1224; NS1225; NS1226; NS1227; NS1228; NS1229; NS1230;  
 CC NS1231; NS1232; NS1233; NS1234; NS1235; NS1236; NS1237;  
 CC NS1238; NS1239; NS1240; NS1241; NS1242; NS1243; NS1244;  
 CC NS1245; NS1246; NS1247; NS1248; NS1249; NS1250; NS1251;  
 CC NS1252; NS1253; NS1254; NS1255; NS1256; NS1257; NS1258;  
 CC NS1259; NS1260; NS1261; NS1262; NS1263; NS1264; NS1265;  
 CC NS1266; NS1267; NS1268; NS1269; NS1270; NS1271; NS1272;  
 CC NS1273; NS1274; NS1275; NS1276; NS1277; NS1278; NS1279;  
 CC NS1280; NS1281; NS1282; NS1283; NS1284; NS1285; NS1286;  
 CC NS1287; NS1288; NS1289; NS1290; NS1291; NS1292; NS1293;  
 CC NS1294; NS1295; NS1296; NS1297; NS1298; NS1299; NS1300;  
 CC NS1301; NS1302; NS1303; NS1304; NS1305; NS1306; NS1307;  
 CC NS1308; NS1309; NS1310; NS1311; NS1312; NS1313; NS1314;  
 CC NS1315; NS1316; NS1317; NS1318; NS1319; NS1320; NS1321;  
 CC NS1322; NS1323; NS1324; NS1325; NS1326; NS1327; NS1328;  
 CC NS1329; NS1330; NS1331; NS1332; NS1333; NS1334; NS1335;  
 CC NS1336; NS1337; NS1338; NS1339; NS1340; NS1341; NS1342;  
 CC NS1343; NS1344; NS1345; NS1346; NS1347; NS1348; NS1349;  
 CC NS1350; NS1351; NS1352; NS1353; NS1354; NS1355; NS1356;  
 CC NS1357; NS1358; NS1359; NS1360; NS1361; NS1362; NS1363;  
 CC NS1364; NS1365; NS1366; NS1367; NS1368; NS1369; NS1370;  
 CC NS1371; NS1372; NS1373; NS1374; NS1375; NS1376; NS1377;  
 CC NS1378; NS1379; NS1380; NS1381; NS1382; NS1383; NS1384;  
 CC NS1385; NS1386; NS1387; NS1388; NS1389; NS1390; NS1391;  
 CC NS1392; NS1393; NS1394; NS1395; NS1396; NS1397; NS1398;  
 CC NS1399; NS1400; NS1401; NS1402; NS1403; NS1404; NS1405;  
 CC NS1406; NS1407; NS1408; NS1409; NS1410; NS1411; NS1412;  
 CC NS1413; NS1414; NS1415; NS1416; NS1417; NS1418; NS1419;  
 CC NS1420; NS1421; NS1422; NS1423; NS1424; NS1425; NS1426;  
 CC NS1427; NS1428; NS1429; NS1430; NS1431; NS1432; NS1433;  
 CC NS1434; NS1435; NS1436; NS1437; NS1438; NS1439; NS1440;  
 CC NS1441; NS1442; NS1443; NS1444; NS1445; NS1446; NS1447;  
 CC NS1448; NS1449; NS1450; NS1451; NS1452; NS1453; NS1454;  
 CC NS1455; NS1456; NS1457; NS1458; NS1459; NS1460; NS1461;  
 CC NS1462; NS1463; NS1464; NS1465; NS1466; NS1467; NS1468;  
 CC NS1469; NS1470; NS1471; NS1472; NS1473; NS1474; NS1475;  
 CC NS1476; NS1477; NS1478; NS1479; NS1480; NS1481; NS1482;  
 CC NS1483; NS1484; NS1485; NS1486; NS1487; NS1488; NS1489;  
 CC NS1490; NS1491; NS1492; NS1493; NS1494; NS1495; NS1496;  
 CC NS1497; NS1498; NS1499; NS1500; NS1501; NS1502; NS1503;  
 CC NS1504; NS1505; NS1506; NS1507; NS1508; NS1509; NS1510;  
 CC NS1511; NS1512; NS1513; NS1514; NS1515; NS1516; NS1517;  
 CC NS1518; NS1519; NS1520; NS1521; NS1522; NS1523; NS1524;  
 CC NS1525; NS1526; NS1527; NS1528; NS1529; NS1530; NS1531;  
 CC NS1532; NS1533; NS1534; NS1535; NS1536; NS1537; NS1538;  
 CC NS1539; NS1540; NS1541; NS1542; NS1543; NS1544; NS1545;  
 CC NS1546; NS1547; NS1548; NS1549; NS1550; NS1551; NS1552;  
 CC NS1553; NS1554; NS1555; NS1556; NS1557; NS1558; NS1559;  
 CC NS1560; NS1561; NS1562; NS1563; NS1564; NS1565; NS1566;  
 CC NS1567; NS1568; NS1569; NS1570; NS1571; NS1572; NS1573;  
 CC NS1574; NS1575; NS1576; NS1577; NS1578; NS1579; NS1580;  
 CC NS1581; NS1582; NS1583; NS1584; NS1585; NS1586; NS1587;  
 CC NS1588; NS1589; NS1590; NS1591; NS1592; NS1593; NS1594;  
 CC NS1595; NS1596; NS1597; NS1598; NS1599; NS1600; NS1601;  
 CC NS1602; NS1603; NS1604; NS1605; NS1606; NS1607; NS1608;  
 CC NS1609; NS1610; NS1611; NS1612; NS1613; NS1614; NS1615;  
 CC NS1616; NS1617; NS1618; NS1619; NS1620; NS1621; NS1622;  
 CC NS1623; NS1624; NS1625; NS1626; NS1627; NS1628; NS1629;  
 CC NS1630; NS1631; NS1632; NS1633; NS1634; NS1635; NS1636;  
 CC NS1637; NS1638; NS1639; NS1640; NS1641; NS1642; NS1643;  
 CC NS1644; NS1645; NS1646; NS1647; NS1648; NS1649; NS1650;  
 CC NS1651; NS1652; NS1653; NS1654; NS1655; NS1656; NS1657;  
 CC NS1658; NS1659; NS1660; NS1661; NS1662; NS1663; NS1664;  
 CC NS1665; NS1666; NS1667; NS1668; NS1669; NS1670; NS1671;  
 CC NS1672; NS1673; NS1674; NS1675; NS1676; NS1677; NS1678;  
 CC NS1679; NS1680; NS1681; NS1682; NS1683; NS1684; NS1685;  
 CC NS1686; NS1687; NS1688; NS1689; NS1690; NS1691; NS1692;  
 CC NS1693; NS1694; NS1695; NS1696; NS1697; NS1698; NS1699;  
 CC NS1700; NS1701; NS1702; NS1703; NS1704; NS1705; NS1706;  
 CC NS1707; NS1708; NS1709; NS1710; NS1711; NS1712; NS1713;  
 CC NS1714; NS1715; NS1716; NS1717; NS1718; NS1719; NS1720;  
 CC NS1721; NS1722; NS1723; NS1724; NS1725; NS1726; NS1727;  
 CC NS1728; NS1729; NS1730; NS1731; NS1732; NS1733; NS1734;  
 CC NS1735; NS1736; NS1737; NS1738; NS1739; NS1740; NS1741;  
 CC NS1742; NS1743; NS1744; NS1745; NS1746; NS1747; NS1748;  
 CC NS1749; NS1750; NS1751; NS1752; NS1753; NS1754; NS1755;  
 CC NS1756; NS1757; NS1758; NS1759; NS1760; NS1761; NS1762;  
 CC NS1763; NS1764; NS1765; NS1766; NS1767; NS1768; NS1769;  
 CC NS1770; NS1771; NS1772; NS1773; NS1774; NS1775; NS1776;  
 CC NS1777; NS1778; NS1779; NS1780; NS1781; NS1782; NS1783;  
 CC NS1784; NS1785; NS1786; NS1787; NS1788; NS1789; NS1790;  
 CC NS1791; NS1792; NS1793; NS1794; NS1795; NS1796; NS1797;  
 CC NS1798; NS1799; NS1800; NS1801; NS1802; NS1803; NS1804;  
 CC NS1805; NS1806; NS1807; NS1808; NS1809; NS1810; NS1811;  
 CC NS1812; NS1813; NS1814; NS1815; NS1816; NS1817; NS1818;  
 CC NS1819; NS1820; NS1821; NS1822; NS1823; NS1824; NS1825;  
 CC NS1826; NS1827; NS1828; NS1829; NS1830; NS1831; NS1832;  
 CC NS1833; NS1834; NS1835; NS1836; NS1837; NS1838; NS1839;  
 CC NS1840; NS1841; NS1842; NS1843; NS1844; NS1845; NS1846;  
 CC NS1847; NS1848; NS1849; NS1850; NS1851; NS1852; NS1853;  
 CC NS1854; NS1855; NS1856; NS1857; NS1858; NS1859; NS1860;  
 CC NS1861; NS1862; NS1863; NS1864; NS1865; NS1866; NS1867;  
 CC NS1868; NS1869; NS1870; NS1871; NS1872; NS1873; NS1874;  
 CC NS1875; NS1876; NS1877; NS1878; NS1879; NS1880; NS1881;  
 CC NS1882; NS1883; NS1884; NS1885; NS1886; NS1887; NS1888;  
 CC NS1889; NS1890; NS1891; NS1892; NS1893; NS1894; NS1895;  
 CC NS1896; NS1897; NS1898; NS1899; NS1900; NS1901; NS1902;  
 CC NS1903; NS1904; NS1905; NS1906; NS1907; NS1908; NS1909;  
 CC NS1910; NS1911; NS1912; NS1913; NS1914; NS1915; NS1916;  
 CC NS1917; NS1918; NS1919; NS19



```

QY      327 CAGCTGCTTAAAG 340
|||||
RESUME 366 standard; DNA; 2407 BP.
AC Q27486
DE 10-FEB-1993 (first entry)
DT GE1 promoter and 5' gene position.
DS expression cassette; root, stem, fertile pollen; as.
OS Oryza sativa.
FH Key Location/Qualifiers
FT Promoter J_2263
FT tata_signal /label= P2L_promoter_region
FT misc_signal /tag= b
FT misc_signal /tag= c
FT cds /label= Transcription_initiation_site
FT /tag= s
FT /label= E1_gene_5'_region
W09311956-A.
PD 20-AUG-1992.
PP 06-FEB-1992. E00274.18
PP 27-SEP-1991. EP-403590.
PP 10-OEC-1991. EP-403352.
PA (PISE ) PLANT GENETIC SYSTEMS NV.
PR WPI 92-1004292181
PR WPI 92-1004292181 F, Mericks S, Scheerlich T;
ST Stamen-specific plant promoters - for producing male-sterile or
male-fertility-restored monocotyledons, e.g. rice
Disclosure: Page 47-48, 59pp English. The promoter regions of stame-
specific rice genes. In These genes were isolated by using male flower-
specific cDNA's as probes (see Q27481-5). The gene sequences isolated
can be used for producing transgenic male-sterile monocots. These
sequences could be used to form expression cassettes which can be
used to provide gene expression in the other parts of the plant, and
do not provide gene expression in the other parts of
the plant that are not involved in the production of fertile pollen.
SQ Sequence 2407 BP; 662 A; 543 C; 507 G; 695 T;
13%; Score 14; DB 4; Length 2407;
Best Match Similarity 100.0%; Fred.No. 3.47e+02;
AC 027486; Conservative 0; Mismatches 0; Indels 0; Gaps 0
DB 338 nucleotide(s) 351

```

```

CC      in the construction of a plant transformation vector comprising a
CC      CC sequence as described which can be used to transform rice
CC      and other plants
CC      CC Sequence 2407 BP;    662 A;   543 G;   507 G;   695 T;
Query Match          13% Score 14; DP 9; Length 2407;
Best Local Similarity 100.0%; Pred. No. 3.47e+02
Matches 14; Conservative 0; Mismatches 0; Indels 0; Caps 0
Db      338 acttetstetctca 351
QY      516 AACCTTTCATCA 529

RESULT 368
ID       T30737 standard; DNA; 2408 BP.
AC       T30737;
DT       20-OCT-1995 (first entry)
PR       Cryptid 1; antibiotic; antimicrobial; defense; inflammation;
KW       antifungal; inflammatory bowel disease; paucitatis;
RW       cancer; tumor; ileitis; dc.
PS       Katus sp.
PE       Location/Qualifiers
FT       exon
           1..1345
           /cgen_start=1374..1176
           /note="exon 1 codes for the 5' untranslated
           region and cryptin-1 prepro sequence"
           1345..1930
           /tag= b
           1930..2408
           /tag= c
           /note="exon 2 codes for cryptin-2 and
           3' untranslated region".
W0616075-A1.
O5-OCT-1995: U13328.
P18-NOV-1994: US-342268.
PPA (SHEI.) SHRINER'S HOSPITAL FOR CRIPPLED CHILDREN.
PRC Ocular Oncology Clinic, University of California at San Diego;
WDI: 96-268527/27.
P-PDSB: R98797.
DR P-PDSB: R98792;
DR P-PDSB: R98791;
DR P-PDSB: R98790;
PPt inflammatory pathologies
PPt activity used paric. In the detection and treatment of
PPt inflammatory diseases.
PPt Claim 31; Page 64-65; 103pp; English.
PPt (R98797) code: US342268 (R98793).
PPt (R98797) code: US342268 (R98793).

```



Tue Nov 17 08:55:26 1998

Tue Nov 17 08:55:26 1998

US-08-007-977-9.PDF

[illegible][illegible]

Tue Nov 17 08:55:26 1998

US-08-887-977-9.129

Page 367

Tue Nov 17 08:55:26 1998

Page 368

[illegible]

ccds\_ for oploid (ant)agonist activity

To isolate opiate receptor genomic clones, 300,000 human genomic clones and a similar number of mouse genomic clones were probed with the 1.1 kb mouse delta opioid receptor clone DOR-1 rat/DNAI isolated. The 3 human clones had very different EcoRI patterns which indicated that three different genes were represented by the human genomic clones which were designated H3, H14 and H20. H4 maps on chromosome 6. It encodes the human kappa opioid receptor. Sequence 2447 bp. 883 A. 438 G. 747 T.

Key Match  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

dbd 2199 ataaagacccaaa 2312  
|||||||  
Cp 647 ATATGAGACCAAA 634

RESEQUIT 372  
ID Q03453 standard; cDNA: 2484 BP.  
OT 20-JUL-1990 (first entry)  
DE cDNA molecule which hybridises to the cDNA of murine IRF-1 encoding a protein with IRF-1 activity.  
PR Interferon regulatory factor-1; interferon-beta gene; cis-elements; es.  
FH Eukaryotic Location/Qualifiers  
cds 154..1200  
//tag= "a  
/product="IRF-1"

EP-355202-A.  
26-FEB-1990.  
24-MOV-1988: 119602 PF  
24-MOV-1988: 119602 PF  
(DNAI/) Tasiguchi T.  
Tasiguchi T.  
NP21: 90-06014/09.  
R2PMB: R03189  
activation of interferon-beta gene transcription by interacting with cis-elements.

Claim 11, Page 41-45; 65pp: English.  
The recombinant molecules induced activation of interferon-beta gene transcription by interaction with cis-elements. The recombinant molecules are used for the production of IRF-1. See also Q03453, -51 and -55; and EP-355190-A. Sequence 2448 BP: 692 A; 630 C; 510 G; 615 T; 1 Others:



```

Query Match      1.38; Score 14; DB 1; Length 2448;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

db 519 aaagaacccaaga 532
|||||
Cp 645 AAGAACCACGA 632

RESRESULT 373
ID ID Q52648 standard; DWA; 2448 BP.
Q52648;
26-MAY-1994 (first entry)
Accession: Q52648;
Keywords: Interferon-beta; with interferon-beta regulating activity.
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

EP-571743.4;
EP-1082-1993.
17-AUG-1989; 115158.
24-AUG-1988; EP-113793.
24-NOV-1988; EP-119602.
Accession: Q52648;
Keywords: Interferon-beta; regulation; gene expression;
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

EP-571743.4;
EP-1082-1993.
17-AUG-1989; 115158.
24-AUG-1988; EP-113793.
24-NOV-1988; EP-119602.
Accession: Q52648;
Keywords: Interferon-beta; regulation; gene expression;
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

P-PDB; R44217;
Accession: P-PDB; R44217;
Keywords: Interferon-beta; regulation; gene expression;
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

Claim 3; Page 36-29; 45pp; English.
The sequence encodes a protein which binds to the repeated oligomer
CC sequence AAGCA and regulatory upstream elements of the human beta
interferon gene. The sequence is identical to the sequence therefore be
used to regulate the expression of interferon-beta.
Accession: Q52648;
Keywords: Interferon-beta; regulation; gene expression;
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

Query Match      1.38; Score 14; DB 9; Length 2448;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

db 519 aaagaacccaaga 532
|||||
Cp 645 AAGAACCACGA 632

RESRESULT 374
ID ID Q130736 standard; DWA; 2457 BP.
Q130736;
23-MAY-1994 (first entry)
Accession: Q130736;
Keywords: Interferon-beta; regulation; gene expression;
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

EP-571743.4;
EP-1082-1993.
17-AUG-1989; 115158.
24-AUG-1988; EP-113793.
24-NOV-1988; EP-119602.
Accession: Q130736;
Keywords: Interferon-beta; regulation; gene expression;
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

P-PDB; R44217;
Accession: P-PDB; R44217;
Keywords: Interferon-beta; regulation; gene expression;
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

Claim 3; Page 36-29; 45pp; English.
The sequence encodes a protein which binds to the repeated oligomer
CC sequence AAGCA and regulatory upstream elements of the human beta
interferon gene. The sequence is identical to the sequence therefore be
used to regulate the expression of interferon-beta.
Accession: Q130736;
Keywords: Interferon-beta; regulation; gene expression;
Regulatory: Interferon-beta; regulation; gene expression;
Interferon element; ss.
DB DB
SS SS Mus musculus. Location/Qualifiers
ccn ccd 154..1050
cds ccd 154..1050
/*tag= a
/*product= Interferon-beta regulatory factor.

```

[illegible]

```

RESULT 375
ID      Q14425 standard; cDNA; 2483 BP.
DESC    Q14425.1997 (figs. entry)
AC      Q14425.1997 (figs. entry)
CD      Q14425.1997 (figs. entry)
DE      TK negative triphosphin, pPRK42.
OS      Tyrosine protein kinase; as.
OS      Mus musculus.
KEYW     Key
FT      1 Location/Qualifiers
FT      2 cda
FT      3
FT      4
FT      5 signal_peptide
FT      6 1..93
FT      7 /tag= a
FT      8
FT      9 mat_peptide
FT      10 93..248
FT      11 /tag= b
FT      12
FT      13 misc_feature
FT      14 1395..2485
FT      15 /product= TK negative trb protein
FT      16
FT      17
FT      18
FT      19
FT      20
FT      21
FT      22
FT      23
FT      24
FT      25
FT      26
FT      27
FT      28
FT      29
FT      30
FT      31
FT      32
FT      33
FT      34
FT      35
FT      36
FT      37
FT      38
FT      39
FT      40
FT      41
FT      42
FT      43
FT      44
FT      45
FT      46
FT      47
FT      48
FT      49
FT      50
FT      51
FT      52
FT      53
FT      54
FT      55
FT      56
FT      57
FT      58
FT      59
FT      60
FT      61
FT      62
FT      63
FT      64
FT      65
FT      66
FT      67
FT      68
FT      69
FT      70
FT      71
FT      72
FT      73
FT      74
FT      75
FT      76
FT      77
FT      78
FT      79
FT      80
FT      81
FT      82
FT      83
FT      84
FT      85
FT      86
FT      87
FT      88
FT      89
FT      90
FT      91
FT      92
FT      93
FT      94
FT      95
FT      96
FT      97
FT      98
FT      99
FT      100
FT      101
FT      102
FT      103
FT      104
FT      105
FT      106
FT      107
FT      108
FT      109
FT      110
FT      111
FT      112
FT      113
FT      114
FT      115
FT      116
FT      117
FT      118
FT      119
FT      120
FT      121
FT      122
FT      123
FT      124
FT      125
FT      126
FT      127
FT      128
FT      129
FT      130
FT      131
FT      132
FT      133
FT      134
FT      135
FT      136
FT      137
FT      138
FT      139
FT      140
FT      141
FT      142
FT      143
FT      144
FT      145
FT      146
FT      147
FT      148
FT      149
FT      150
FT      151
FT      152
FT      153
FT      154
FT      155
FT      156
FT      157
FT      158
FT      159
FT      160
FT      161
FT      162
FT      163
FT      164
FT      165
FT      166
FT      167
FT      168
FT      169
FT      170
FT      171
FT      172
FT      173
FT      174
FT      175
FT      176
FT      177
FT      178
FT      179
FT      180
FT      181
FT      182
FT      183
FT      184
FT      185
FT      186
FT      187
FT      188
FT      189
FT      190
FT      191
FT      192
FT      193
FT      194
FT      195
FT      196
FT      197
FT      198
FT      199
FT      200
FT      201
FT      202
FT      203
FT      204
FT      205
FT      206
FT      207
FT      208
FT      209
FT      210
FT      211
FT      212
FT      213
FT      214
FT      215
FT      216
FT      217
FT      218
FT      219
FT      220
FT      221
FT      222
FT      223
FT      224
FT      225
FT      226
FT      227
FT      228
FT      229
FT      230
FT      231
FT      232
FT      233
FT      234
FT      235
FT      236
FT      237
FT      238
FT      239
FT      240
FT      241
FT      242
FT      243
FT      244
FT      245
FT      246
FT      247
FT      248
FT      249
FT      250
FT      251
FT      252
FT      253
FT      254
FT      255
FT      256
FT      257
FT      258
FT      259
FT      260
FT      261
FT      262
FT      263
FT      264
FT      265
FT      266
FT      267
FT      268
FT      269
FT      270
FT      271
FT      272
FT      273
FT      274
FT      275
FT      276
FT      277
FT      278
FT      279
FT      280
FT      281
FT      282
FT      283
FT      284
FT      285
FT      286
FT      287
FT      288
FT      289
FT      290
FT      291
FT      292
FT      293
FT      294
FT      295
FT      296
FT      297
FT      298
FT      299
FT      300
FT      301
FT      302
FT      303
FT      304
FT      305
FT      306
FT      307
FT      308
FT      309
FT      310
FT      311
FT      312
FT      313
FT      314
FT      315
FT      316
FT      317
FT      318
FT      319
FT      320
FT      321
FT      322
FT      323
FT      324
FT      325
FT      326
FT      327
FT      328
FT      329
FT      330
FT      331
FT      332
FT      333
FT      334
FT      335
FT      336
FT      337
FT      338
FT      339
FT      340
FT      341
FT      342
FT      343
FT      344
FT      345
FT      346
FT      347
FT      348
FT      349
FT      350
FT      351
FT      352
FT      353
FT      354
FT      355
FT      356
FT      357
FT      358
FT      359
FT      360
FT      361
FT      362
FT      363
FT      364
FT      365
FT      366
FT      367
FT      368
FT      369
FT      370
FT      371
FT      372
FT      373
FT      374
FT      375
FT      376
FT      377
FT      378
FT      379
FT      380
FT      381
FT      382
FT      383
FT      384
FT      385
FT      386
FT      387
FT      388
FT      389
FT      390
FT      391
FT      392
FT      393
FT      394
FT      395
FT      396
FT      397
FT      398
FT      399
FT      400
FT      401
FT      402
FT      403
FT      404
FT      405
FT      406
FT      407
FT      408
FT      409
FT      410
FT      411
FT      412
FT      413
FT      414
FT      415
FT      416
FT      417
FT      418
FT      419
FT      420
FT      421
FT      422
FT      423
FT      424
FT      425
FT      426
FT      427
FT      428
FT      429
FT      430
FT      431
FT      432
FT      433
FT      434
FT      435
FT      436
FT      437
FT      438
FT      439
FT      440
FT      441
FT      442
FT      443
FT      444
FT      445
FT      446
FT      447
FT      448
FT      449
FT      450
FT      451
FT      452
FT      453
FT      454
FT      455
FT      456
FT      457
FT      458
FT      459
FT      460
FT      461
FT      462
FT      463
FT      464
FT      465
FT      466
FT      467
FT      468
FT      469
FT      470
FT      471
FT      472
FT      473
FT      474
FT      475
FT      476
FT      477
FT      478
FT      479
FT      480
FT      481
FT      482
FT      483
FT      484
FT      485
FT      486
FT      487
FT      488
FT      489
FT      490
FT      491
FT      492
FT      493
FT      494
FT      495
FT      496
FT      497
FT      498
FT      499
FT      500
FT      501
FT      502
FT      503
FT      504
FT      505
FT      506
FT      507
FT      508
FT      509
FT      510
FT      511
FT      512
FT      513
FT      514
FT      515
FT      516
FT      517
FT      518
FT      519
FT      520
FT      521
FT      522
FT      523
FT      524
FT      525
FT      526
FT      527
FT      528
FT      529
FT      530
FT      531
FT      532
FT      533
FT      534
FT      535
FT      536
FT      537
FT      538
FT      539
FT      540
FT      541
FT      542
FT      543
FT      544
FT      545
FT      546
FT      547
FT      548
FT      549
FT      550
FT      551
FT      552
FT      553
FT      554
FT      555
FT      556
FT      557
FT      558
FT      559
FT      560
FT      561
FT      562
FT      563
FT      564
FT      565
FT      566
FT      567
FT      568
FT      569
FT      570
FT      571
FT      572
FT      573
FT      574
FT      57
```

Query Match Local Similarity 100.0%: Score 14; DP 36; Length 2489;  
Matches 14; Conservative 0; Mismatches 0; Indels 0;

DD 1460 gctgtggttgatt 1473  
| | | | | | | | | |  
Seq 745 GCTGTTGGTGTACT 758

SEQUENCE 376  
756-775 standard; cDNA; 2503 BP.  
AC TGA6715;  
AT CAC  
CGT 26-FEB-1998 (first entry)  
DE MSRV-1 clone LTRMGAG1.  
NEW amplification product from tissue culture supernatant of  
MDM cells infected with rhumatoid arthritis, derived from  
MO9706280-A1.  
PM MO9706280-A1.  
PN CM  
PS 03-FEB-1997; P01344  
PT 03-FEB-1997  
PR 03-OCT-1995; FR-009643.  
PP (JNMR) B10 MEMRIEX.  
RA Bedin F, Berene F, Jolivet-Reynaud C, Komurian-Pradel F,  
Roux S, Virelizier JL, Monod H, Dubouilh G, Perron B:  
MSRV-1 viral DNA and nucleotide fragments associated with mul-  
tiple sclerosis and rheumatoid arthritis - also related peptide(s) as  
antibodies, used for diagnosis, treatment and as vaccines  
sequences 796468-779 represent claimed nucleic acid sequences de-  
rived from the multiple sclerosis related virus 1 (MSRV-1), especially  
the env, pol and part of the gag gene region, but exclude the re-  
presented by T96527. This sequence is from clone LTRMGAG1 at  
position 796527.  
CC Cells expressing the MSRV-1 retrovirus in a patient with multi-  
ple sclerosis. The sequence was isolated by PCR using primers T96527  
and T96528.  
CC The above sequences are associated with multiple sclerosis and  
from this virus were associated with rheumatoid arthritis. Nucleic acid se-  
quences were used in the diagnosis of multiple sclerosis  
and rheumatoid arthritis.  
CQ Sequence 2503 BP. 631 C; 563 G; 631 T;  
GC

Query Match Local Similarity 100.0%: Score 14; DP 36; Length 2503;  
Matches 14; Conservative 0; Mismatches 0; Indels 0;

DD 2188 tgaactggctttcct 2201  
| | | | | | | | | |  
Seq 136 TGAAGTCGCTTCT 149



ID	777	
ISD	N60895 standard; DNA;	2546 BP.
AC	N60895;	
DT	16-OCT-1991	(first entry)
DE	Fragment pHS 53 homologous to human chromosome 11 pre-central region.	
DI	Kretzschmar; as	
DI	Kretzschmar; as	
PN	SU1201108-A	
PD	07-JAN-1985	
PP	02-AUG-1984;	778107.
PR	02-AUG-1984;	SU-778107.
PT	(Gibco) BRL PHOTACH HEALTH.	
QA	Gladstein, NM; review 15,	Takovlev AG;
DR	WPI:	86-2112013/32
PT	Marking of eleventh human chromosome - involves using recombinant plasmid DNA pHs 53 and genome DNA pHs 53 fragment.	
PT	Claim 1: Page 1: 9pp; Russian.	
CC	The probe may be tritium labelled and used to mark the pre-central region of human chromosome 11, useful in prenatal diagnosis,	
CC	pathology etc.	
SC	Sequence	2546 BP: 743 A: 493 C: 525 G: 795 T:

```

Query Match      1.3%      Score 14: Db 2: Length 2546:
Best Local Similarity 100.0%      pred No 3: 4e+02:
Matches 14: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 594 consnaactcatt 607
QY 941 TCAGAAACTTATT 954

```

RESULT	378
ID	G64889 standard; cDNA; 2546 bp.
AC	G64889;
CC	
DT	07-NOV-1994 (first entry)
ED	Insertion phase in insert pDS16.
KW	Sermyel protein transferase; mutation; framesylation; ss.
OR	
OS	Homo sapiens.
FH	Key
FT	Location/Qualifiers
FT	cds
FT	23..1345
FT	/note= "alpha subunit"
FT	1348..2487
FT	/note= "beta subunit"
FT	cds
FT	23..1345
FT	/note= "alpha subunit"
FT	1348..2487
FT	/note= "beta subunit"
PN	W094I0184-A.
PD	11-MAY-1994.
PF	29-OCT-1993; U10442.
PR	U10442.
RN	(MRT)-INRECI JY.
RI	Diedl AE, Gibbs JR, Kohl NE,
PI	Omer CA;

Mon Nov 17 08:55:26 1998

DR WP1: 96-268927/27.  
DR P-PDSB: R98788.  
DR P-PDSB: R98793.  
DR P-New isolated crypidin peptide(s) - which have antimicrobial  
DR activity, used partic. in the detection and treatment of  
DR  
PS Claim 31: Page 63-66; 103pp: English.  
CC The rat crypidin-3 gene (T30738) codes for the precursor  
CC (R98788) of crypidin-2 (R98793), an antimicrobial peptide  
CC that exhibits activity against a broad range of intestinal  
CC and other pathogens. The crypidin-3 gene was isolated from  
CC (see also T30736-37) were isolated from a Sprague-Dawley  
CC genomic library cloned in EMBL3. The crypidin-3 gene, or  
CC corresponding cDNA (see also T30738), can be used to produce  
CC the protein for use in treating inflammatory  
CC pathologies of the intestine.  
SQ Sequence 2551 BP: 689 A; 543 C; 563 G; 756 T;  
Query Match 138; Score 15; DB 27; Length 2551;  
Best Local Similarity 100.0% P-Id M: 87-061.  
Matches 15: Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1153 gagctctcaaggagat 1167  
Cc 560 gagctctcaaggagat 646

RESULT	380	standard; cDNA; 2568 bp.
AC	72923	
AC	T62842;	
DT	08-JUN-1997	(first entry)
DT	Rat oestrogen receptor beta (ER-beta) cDNA.	
KW	Oestrogen receptor; oestrogen receptor; ER-beta;	
KW	Orphan receptor; progesterone receptor; Progesterone	
KW	ovary cancer; endometrial cancer; Benign prostatic hyperplasia;	
KW	ovary cancer; endometrial cancer; Endometriosis; Osteoporosis;	
KW	environment; Pollutant; ss.	
OS	Rattus sp.	
PH	Key	Location/Qualifiers
PR	cds	744..1881
PT		/note= "alternative start codon at 448...450"
FT		
PD	W09709348-A2.	
PD	13-MAR-1997.	E03033
PR	SEB-1986:	GB-003572.
PR	SEB-1986:	GB-003572.
PR	15-MAY-1986:	GB-003550.
PR	11-APR-1986:	GB-007532.
PR	08-MAY-1986:	GB-008576.
PR	(KARO.) KMO BIC AB.	
PT		
DT	NFL-97-192842/1	Ingram J, Kuiper GG;

DR WP1: 94-167373/20.  
DR #PBUB: R5482; anti-protein transferase probe - used in assays  
DR human placenta cDNA library in lambda still using a  
DR for cpds with inhibitory activity for the identification of  
PT anticancer agents  
PT Disclosure: Fig 5; 69pp; English.  
CC The cDNA encoding the human alpha and beta subunits of Pfrase was  
CC derived from a human placenta cDNA library in lambda still using a  
CC bovine Pfrase gene as a template. The protein encoded by the  
CC inhibitory activity of a cpd. In the transfection of a protein  
CC substrate. The assay can be used to identify anticancer agents.  
CC See also Q64896-8.  
SQ Sequence 2546 BP; 589 A; 607 C; 653 G; 597 T;  
SQ Query Match 1.3%; Score 15; DB 10; Length 2546;  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 307 tctgattcgagccggc 321  
CP | ||||| |||||  
CP 450 tttgatcgacccgc 436

RESULT 372  
 AC 10718 standard; DNA: 2551 BP.  
 AC 10718 AC  
 T10718 20-OCT-1996 (first entry)  
 DE Rat cryptid 3 gene.  
 DE Cryptid 3; antitoxic; antimicrobial;  
 DE Cryptid 3; antitoxic; defensin; inflammation;  
 DE antinflammatory; inflammatory bowel disease;  
 DE cancer; tumour; ileitis; da.

Accession	Location/Qualifiers
U08	1..1482
U09	1..1482
U10	1..1482
U11	1..1482
U12	1..1482
U13	1..1482
U14	1..1482
U15	1..1482
U16	1..1482
U17	1..1482
U18	1..1482
U19	1..1482
U20	1..1482
U21	1..1482
U22	1..1482
U23	1..1482
U24	1..1482
U25	1..1482
U26	1..1482
U27	1..1482
U28	1..1482
U29	1..1482
U30	1..1482
U31	1..1482
U32	1..1482
U33	1..1482
U34	1..1482
U35	1..1482
U36	1..1482
U37	1..1482
U38	1..1482
U39	1..1482
U40	1..1482
U41	1..1482
U42	1..1482
U43	1..1482
U44	1..1482
U45	1..1482
U46	1..1482
U47	1..1482
U48	1..1482
U49	1..1482
U50	1..1482
U51	1..1482
U52	1..1482
U53	1..1482
U54	1..1482
U55	1..1482
U56	1..1482
U57	1..1482
U58	1..1482
U59	1..1482
U60	1..1482
U61	1..1482
U62	1..1482
U63	1..1482
U64	1..1482
U65	1..1482
U66	1..1482
U67	1..1482
U68	1..1482
U69	1..1482
U70	1..1482
U71	1..1482
U72	1..1482
U73	1..1482
U74	1..1482
U75	1..1482
U76	1..1482
U77	1..1482
U78	1..1482
U79	1..1482
U80	1..1482
U81	1..1482
U82	1..1482
U83	1..1482
U84	1..1482
U85	1..1482
U86	1..1482
U87	1..1482
U88	1..1482
U89	1..1482
U90	1..1482
U91	1..1482
U92	1..1482
U93	1..1482
U94	1..1482
U95	1..1482
U96	1..1482
U97	1..1482
U98	1..1482
U99	1..1482
U100	1..1482

119-08-887-977-9

P-PCUB: W14723.  
The isolated oestrogen receptor beta - used to develop prods. for treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular disease.

CC Claim 6: Fig 1: 45pp: English.  
CC A cDNA clone (7828d) codes for a novel rat oestrogen receptor-  
CC protein. The deduced amino acid sequence of the protein contains  
CC degenerate primers based on conserved sequences of the DNA-  
CC binding domain and ligand binding domain of nuclear receptors were  
CC used to amplify rat prostate cDNA. The PCR prod. was used to  
CC screen a rat prostate cDNA library to obtain the ER-beta cDNA  
CC (from a human testis cDNA library). Rat, human and mouse ER-beta  
CC from (71721-25) can be used to locate molecules for use in the  
CC treatment of cardiovascular diseases, central nervous system  
CC disorders, cancer, osteoporosis, osteoarthritis, osteopenia,  
CC breast hyperplasia and to test environmental chemicals for  
CC oestrogenic activity.

CC Sequence 2569 BP: 634 A: 654 G: 617 F:  
CC

Query Match 13%; Score 15; DB 29; Length 2569;  
Best Local similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CP	267	ACTAAGCAGCAGAG	253
		RESULT	301
		ID	Q6237 standard; DNA: 2568 BP.
		ID	Q6237; AC
		ID	28-JAN-1991 (first entry)
		ID	Sequence encoding cellulase.
		ID	Enzyme: aa.
		ID	Source: Mytilus IPO 3288
		ID	Key
		ID	Location/Qualifiers
		ID	819..1859
		ID	/tag= a
		ID	/product=cellulase
		ID	302722698-A
		ID	05-SEP-1990.
		ID	27-FEB-1989: 045570.
		ID	27-FEB-1989: JP-045570.
		ID	(EPR-) SHIN EIKIET SAKETO.
		ID	Accession: J01421
		ID	WPI: 90-21550/42.
		ID	P-980B; R07356.
		ID	Cellulase coding DNA sequence - can produce cellulase by
		ID	bio:engineering method



[illegible]

```

CC the 85kd polypeptide fragments and a 70kd polypeptide, diagnostic reagent
CC for Chagas' disease can be made. 85kd polypeptide and 70kd polypeptide
CCCC are recognised by all of the Chagasic sera tested, but not by Leishmania
CC sera.
Query Match      1.3%   Score 14; DB 1: Length 2572;
Beat-Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Seq. Sequence    2572 BP; 629 A; 623 C; 794 G; 526 T;

db 1000 gtgaagcgcacgpc 1013
Cp 1991 GTGAGGACGACGCC 1078

RESULT 383
ID ID Q1385 standard; cDNA; 2607 BP.
DC DC Q1385-1993 (first entry)
DE DE Plasium sativum L PAL gene.
DB DB Phenylalanine ammonia lyase: phenyl propanoid isoflavonoid; plant;
KW KW pea. es.
NC NC Key word sativum. Location/Qualifiers
FT FT cds 126..2456 /tag= a
FF FF 104310295.A.
PR PR 19 APR-1991; 097897.
PP PP 25-APR-1991; 097897.
PP PP 26-APR-1991; JP-097697.
PA PA (XMS ) TAKASAGO PHENOLIC CO LTD.
PA PA P-FEMS: R30077/01.
PP PP New phenylalanine ammonia lyase gene from plasium sativum L -
PP PP plasmid contg. gene, and E. coli transformed with plasmid. for
PP PP Diseno of lysae 7; 9pp; Japanese.
PS PS The sequence is that of the phenylalanine ammonia lyase (PAL) gene
PS PS from plasium sativum L. The gene may be used to produce PAL
CC recombinantly. Plants contg. the PAL gene can contain a large amt.
CC of isoflavonoid.
CC See also Q30101.
CC Sequence 2607 BP; 795 A; 485 C; 577 G; 750 T;
CC Query Match      1.3%   Score 15; DB 6: Length 2607;
CC Beat-Local Similarity 100.0%; Pred. No. 4.0e+02;
CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC Seq. Sequence    2607 BP; 795 A; 485 C; 577 G; 750 T;

db 1992 chaqtcgtctaaag 1946
db 1993 chaqtctttaaagg 341

```

[illegible][illegible]







KW	Rice:	starch, transit peptide;	pectin; cereal, amiopectin; seeds;
KW	reverse	transcriptase; plaque; ss.	
OS	Oryza	ativa	
PH	Key	Location/Qualifiers	
FT	5'utr	1..17	
FT	1	1..17	a
FT	mrna	1..2741	
FT	transit_peptide	/.*tag= b	
FT	18..197	18..197	c
FT	mat_peptide	198..2477	
FT	1	198..2477	d
FT	3'utr	2478..2733	
FT	1	2478..2733	e
PN	J05317/057-A		
PD	03-DEC-1993		
PD	30-MAR-1994	102499	
PR	20-SEP-1991	JP-268617	
PR	01-01-1991	035A1 SHOKUKUTU BIO KENTUSHO KK.	
WP1:	94-011022/01		
DR	P-PS09	R47468	
PT	Gene	CDNA for rice starch branching enzyme for varied amino	
PT	protein in cereal	starches structural gene specified by basic	
PT	fragment	originated from rice genome contig. gene	
PS	Claim 1:	Page 57: 21pp; Japanese.	
CC	The	sequence shows a cDNA encoding a branching enzyme of rice	
CC	starch in	cereal particles by introducing this basic sequenc into	
CC	a rice	plant. This process can be used to improve the taste of the	
CC	rice.		
SC	Sequence	2733 BP: 752 A; 558 C; 674 G; 749 T;	
Query Match	1.3%;	Score 14; DB 10; Length 2733;	
Best Local Similarity	100.0%;	Pred. No. 3.47e+02;	
Matches	14;	Conservative 0; Mismatches 0; Indels 0; Gaps	
Db	1515	aaagcacaagcgga 1528	
Cp	264	AAGACAAAGACGA 251	
RESULT	392		
ID	G62135	standard; cDNA to mRNA; 2733 BP.	
CC	06-MAR-1995	(first entry)	
DE	KW	Rice starch branching enzyme coding sequence.	
DE	KW	Rice starch branching enzyme; oryza ativa; amylopectin; albulmen;	
DE	starch; ss.		
DE	Oryza	ativa.	
PH	Key	Location/Qualifiers	

```

FT      misc_difference 233...239
FT      /note= "bases n at positions 233-240 are unclear
FT      in the specification"
FT
FN      CA108606-A.
FP      15-OCT-1993; 108606
FP      15-OCT-1993; CA-108606.
PR      (INRM ) INST NAT SANTE & RECH MEDICALE.
PR      Aumont J, Mandel J, Mosser J, Sarde C:
PI      Oncogene 71:2247-2253.
PI      P-SBDB: R76110.
DR      New nucleic acid responsible for adreno-leuco-dystrophy - related
FT      probes, proteins and antibodies, useful for diagnosis and treatment
FT      of adrenoleucodystrophy. English, the red pigment gene of an ANK
FT      protein was used to isolate clones from an Xg28 cosmid library.
CC      Following hybridizations and nested PCR, a HeLa cell cDNA library
CC      was screened, and a cDNA fragment contg. the complete coding
CC      sequence of the complementary DNA of the complementary cDNA of
CC      hematopoietic cells with the sequence in vivo will allow therapy
CC      of ALD or ANK.
CQ      Sequence 2746 BP: 439 A: 964 C: 815 G: 466 T:
Query Match 133: Score 14; DB 15; Length 2746;
Best Local Similarity 100.0%; Pred.no 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps:
Db          640 tctctgcggcgag 653
            |||||
OY          1012 TCCTGTCCGGGAG 1025

RESULT 394
RESLT3315 standard; cDNA: 2750 BP.
AC      TR3935;
AC      TR3935; 1997 (first entry) ALD; wild-type gene.
DT      E-REB-1997; 1997
KW      Adrenoleucodystrophy; gene therapy; retroviral vector; M48;
KW      adrenoleucopathy; membrane protein; long chain fatty acid oxidation
ds       ds
Key      Key
EP      Key leucoplasms.
FT      Location/Qualifiers
FT      cdscd      387..2624
FT      /tag=a
FT      /product=membrane_protein
FT      /size=189
FT      /tag=b
FT      /label=Exkl
FT      /note= "Exkl corresponds to this region"
FT      ..ccccc..

```



PD 16-JUL-1996. PRO058.  
 PR 13-JAN-1995. PR-00376.  
 PI (INRM ) INST NAT SANTE & RECH MEDICALE.  
 PI Aubourg P, Mandel JL, Mosser J, Sarde CO;  
 DR WPI: 96-34286/34.  
 PT Recombinant viral vector contg. DNA for correcting  
 PT adrenoleuco-dystrophy - and immunologically tolerable cells contg.  
 PT this vector, useful in gene therapy  
 CC A family of related DNA insertion sequences (ISM) in the genome of  
 CC (adrenoleucodystrophy) cDNA derived from the present sequence was  
 CC inserted into retroviral vector M48 so that it was under control  
 CC of the mouse phosphotyrosine kinase (PTK) promoter. The resulting  
 CC vector was cell infected with vector PTK-neomycin into amphotropic  
 CC cells which were also positive with anti-AD antibodies were incubated  
 CC with skin fibroblasts of an ALD patient. After two rounds of infection,  
 CC about 70% of cells were expressing normal ALD protein and integration  
 CC of the ALD gene confirmed by Southern blotting.  
 CC Sequence 2750 BP; 491 A; 513 G; 885 G; 455 T;  
 SQ

Query Match 1.33; Score 14; DB 25; Length 2750;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 643 tctctgctggag 656  
 QY 1012 TCTCTGCTGGAG 1025

RESULT 395  
 ID N80976 standard; DNA: 2762 BP.  
 DC 12-SEP-1990 (first entry)  
 DE Sequence of 2 contiguous lengths of DNA (contigs C1 and C2) on pMB22  
 DE contg. Mycobacterial insertion sequences (ISM)  
 KW Mycobacterial insertion sequences; Mycobacterium; pMB22; NCIB 12461;  
 KW Mycobacterial disease prevention; vaccine;  
 OS Mycobacterium.  
 FH Key Location/Qualifiers  
 FT misc\_feature 1..1280  
 FT Note: a  
 FT /note="contig C1"  
 FT misc\_feature 1261..2762  
 FT /tag= b  
 FT /note="contig C2"  
 EP-288306-A.  
 PD 26-OCT-1988.

PF 22-APR-1988. 303657.  
 PR 24-APR-1987. GB-009803.  
 PI (MCPA/) McFadden J.  
 PI C. J. McFadden, Raymond-Taylor J;  
 DR WPI: 88-30136/43.  
 PT DNA insertion sequences of mycobacterial origin -  
 PT producing proteins for vaccines, immuno therapy or diagnostics  
 CC A family of related DNA insertion sequences (ISM) in the genome of  
 CC disease isolates of mycobacteria has been isolated. The first example of  
 CC ISM (ISM-1) was identified and characterised in the clone pMB22 (NCIB  
 CC 12461) derived from Mycobacterium tuberculosis H37Rv. Mycobacterial isolate  
 CC from Crohn's disease. The ISM-1 was also identified in Mycobacterium  
 CC paratuberculosis cultured from cattle and other animals suffering from  
 CC Johne's disease. N80976 gives the sequence of the portion of pMB22  
 CC which spans the ISM-1 insertion element. The DNA can be used as DNA or  
 CC probe to identify the ISM-1 in other strains of Mycobacterium. The ISM-1  
 CC mycobacteria in environmental and disease samples and in vitro  
 CC closely related mycobacterial strains. ISM derived proteins or peptides  
 CC can be used in diagnostics or for producing antibodies or as vaccines  
 CC Sequence 2762 BP; 432 A; 897 G; 918 G; 513 T; 2 Others;  
 SQ

Query Match 1.33; Score 14; DB 1; Length 2762;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 2412 tgaattggggctt 2425  
 QY 611 TGAATTGGGGCTT 624

RESULT 396  
 ID N80976 standard; DNA: 2763 BP.  
 DC 12-SEP-1990 (first entry)  
 DE Human disintegrin metalloprotease (aggrinase) gene.  
 DE Disintegrin; metalloprotease; arthritis; osteoarthritis; diagnosis;  
 KW aggrinase; arthropathy; osteoporosis; ankylosing spondylitis;  
 KW Human disease; human; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 17..2416  
 FT Note: a  
 FT /tag= a  
 PD 04-SEP-1987.  
 PF 28-FEB-1997; U03217.  
 PR 01-MAR-1996; US-012679.  
 PR (PROC ) PROCTER & GAMBLE CO.

PA (UYCA-) UNIV CASE WESTERN RESERVE.  
 PI Haqit T, Tindal MH;  
 PR 19-SEP-97 44826/41.  
 PT New human disintegrin metalloprotease - is differentially expressed  
 PT in arthritis, used to screen for drugs and to diagnose  
 PT metalloprotease-mediated diseases, e.g. osteoarthritis  
 CC The protein encoded by this gene is a human disintegrin  
 CC metalloprotease (see #3233). It designated aggrinase that is  
 CC up-regulated in human chondrocytes during osteoarthritic conditions  
 CC and which is involved in tissue remodeling and breakdown. It was  
 CC obtained by screening human cDNA libraries with a partial cDNA  
 CC stimulated human articular (femoral head) chondrocytes using  
 CC differential display analysis. Disintegrin nucleic acids can be  
 CC used in the construction of recombinant expression vectors for  
 CC a marker to monitor the development of osteoarthritis. The gene provides  
 CC asymptomatic stages, and the progression of osteoarthritis.  
 CC Inhibitors of the disintegrin, including antisense oligonucleotides,  
 CC can potentially be used in the treatment of metalloprotease-mediated  
 CC diseases such as osteoarthritis, ankylosing spondylitis, rheumatoid or inflammatory arthritis, joint and Lyme  
 CC disease.  
 CC Sequence 2763 BP; 849 A; 537 C; 610 G; 767 T;  
 SQ

Query Match 1.33; Score 14; DB 36; Length 2763;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 2487 agtctatctctca 2500  
 QY 498 AGTCATCTCTCA 511

RESULT 397  
 ID Q13239 standard; cDNA: 2781 BP.  
 AC Q13239;  
 DT 29-OCT-1991 (first entry)  
 KW Heat shock factor; sequence.  
 OS Drosophila.  
 FH Key Location/Qualifiers  
 FT cds 229..2304  
 FT Note: a  
 FT /tag= a  
 FT poly\_a\_signal 2729..2727  
 FT /tag= b  
 FT poly\_a\_site 1757..1781  
 FT /tag= c  
 PD 087617901-A.

PD 16-JUL-1991. 617901.  
 PR 26-NOV-1990; US-617901.  
 PI (USSH ) NAT INST OF HEALTH  
 PI C. J. McFadden, Raymond-Taylor J, Rabindran S;  
 DR WPI: 91-25234/74.  
 DR P-PDS: R13502.  
 PT DNA encoding Drosophila and human heat shock factor proteins -  
 PT used for developing prods. for studying stress and disease states -  
 CC The sequence encodes Drosophila heat shock factor protein and was  
 CC obtained by screening a Drosophila genomic library with oligo-  
 CC nucleotide probes (Q13237, Q13238) based on the HSF amino acid  
 CC sequence. The protein encoded by this gene is a Drosophila heat shock  
 CC factor protein. It is a transcription factor that binds to DNA in  
 CC other organisms and also for the detection of stress or a diseased  
 CC state in living systems. The gene can be used to increase  
 CC expression of other gene prods. by cotransfecting the HSF gene  
 CC into cells with other genes linked to heat shock elements. It can be  
 CC linked to other genes linked to heat shock elements and  
 CC introduced into transgenic mice as a tool for eliciting increased  
 CC or chronic stress response conditions as a model for how tissues  
 CC respond to chronic stress conditions such as those caused by viral  
 CC infection, chemical or mechanical stress. See also Q13240 and  
 CC Q13241.  
 CC Sequence 2781 BP; 831 A; 631 C; 690 G; 629 T;  
 SQ

Query Match 1.33; Score 14; DB 2; Length 2781;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 990 gaccagtgagaccy 1003  
 QY 1050 GACCAGTGAGACC 1063

RESULT 398  
 ID N80976 standard; DNA: 2800 BP.  
 AC Q5561;  
 DT 15-JUL-1994 (first entry)  
 DE Meningitis M2394 transferrin receptor Tbp1 subunit DNA sequence.  
 KW Transferrin receptor; Tbp1 subunit; strain M2394; meningitis vaccine;  
 KW Meningitis; vaccine; strain M2394.  
 OS Neisseria meningitidis (strain M2394).  
 FH Key Location/Qualifiers  
 FT cds 40..2766  
 FT Note: a  
 FT /tag= a  
 FT signal\_peptide 40..111  
 FT /tag= b  
 FT mat\_peptide 112..2763  
 FT /tag= c







```

Db      305 aaagatcttaaga 318
QY      958 MAGATCTTGAAGA 971

RESULT 403
ID      Q87441 standard; cDNA; 2854 bp.
AC      AC   Q87441; 1995
GI      GI   Z146012; 1995 (cflret entry) RNA.
KW      KW     Semaphorin; grasshopper; human; vaccinia virus; Drosoephila; Tribolium;
        KW     varicella major virus; smallpox; human; semaphorin receptor binding activity;
        KW     modulation; nerve cell growth; immune response; viral pathogenesis;
        KW     neurological disease; neuro-regeneration; oncological infection; ds.
        KW     Grasshopper sp. Location/Qualifiers
FF      FF     CDS
FT      FT     431..2643
TT      TT     /tag= a
          /product= grasshopper semaphorin
W09507706-A
23-MAR-1995
PP      PP     13-SEP-1994; U10131.
PR      PR     13-SEP-1993; GS-121713.
RT      RT     BATTLEY DR, CALIFORNIA.
PI      PI     BATTLEY DR, Goodman CS, Kolodkin AL, Matthes D;
          WPI: 95-131177-17.
DR      DR     P-9508: R1737.
CC      CC     The grasshopper semaphorin peptide(s) and polypeptide(s) - are
          potential modulators of neuronal development and regeneration
          Example 1: Page 68-72; 101pp; English.
          The sequence of the cDNA encoding the grasshopper semaphorin I protein.
          CC IT Proteins encoded by the grasshopper semaphorin I, human semaphorin
          CC IT semaphorin I, mouse semaphorin I, Tribolium semaphorin I (Q87441), or
          CC semaphorin I and II (Q87444-5). Tribolium semaphorin I (Q87441) or
          CC generate a series of peptides (R70370-R70418), which retain semaphorin
          CC receptor binding activity. The semaphorin derived or semaphorin receptor
          CC responsible for neurite outgrowth, axon guidance, synapse formation, and
          CC responsiveness and viral pathogenesis. They can be used in diagnosis and
          CC treatment of neurological disease and neuro-regeneration, immune
          CC modulation and diagnosis and treatment of viral and oncological infection
          CC and diseases.
          Sequence 2854 BP; 804 A; 637 C; 650 G; 763 T;
          Query Match
          Best Local Similarity 100.0%; Pred. No. 9.47e+02;
          Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
          Db      366 tggcttcactgcagg 379

```

QY
912 TGGCTGACACTT G25
RESULT 404
ID Q78280 standard; DNA: 2857 BP.
CD CDS
OS NCN-1995
OT 06-JUN-1995 (first entry)
T 2-nivem GAPDH spdh gene.
KW Alanine racemase; glyceraldehyde-3-phosphate-dehydrogenase;
KX GAPDH; cytoplasmic; immunosuppressive; promoter; da.
YC Key:
FH Cyclopladim n.nucleotide
FT exon 512..549 Location/Qualifiers
TT exon /*tag= a
TT intron 550..664 b
TT /label= Intron-1
TT exon 665..801 c
TT /*tag= c
TT intron 802..1218 d
TT /label= Intron-2
TT exon 1219..1331 e
TT /*tag= e
WT W99J26506-A.
PV 10-MAY-1994.
PP 23-APR-1994; E01272.
PR 23-APR-1993; DE-312856.
PR 04-MAY-1993; DE-314616.
PA S(NAO) SANDOZ PATENT GNTH.
PA (SANO) SANDOZ PATENT GNTH.
PA (SANO) SANDOZ-ERFINDUNGEN VERW GES MBH.
PI Socher HP, Schneider-scherrer Z, Schoeberdorfer K;
DQ WI94_358281/44.
DG New nucleic acid encoding euukaryotic alanine racemase - and
DT related vectors, host cells, and recombinant enzyme, useful for
PT producing cyclopurin derivs. or increasing cyclopurin prodn.,
PS Claim 6: Page 51-52; 8pp; German.
PC A cosmid bank of 2-nivem ATCC 34921 genomic DNA was screened with
CC a fragment of the glyceraldehyde-3-phosphate-dehydrogenase (GAPDH)
CC gene from Escherichia coli chromosome. NCI 8.2 kb fragment was 665 bp
CG identical to F0987711 chimera construct.
CH HindIII-HindII fragment of this was used to screen cosmid DNA to
CC identify A(nu) fragment (F078278) containing the GAPDH promoter.
CC A sequence (F078278) encoding the GAPDH gene (spdh) was isolated from
CC Saccharomyces cerevisiae strain Y28780 as sequence of the genomic spdh structural
CC gene.

CC	R55965	The CARP8 promoter is useful for expressing foreign genes in T. niivemus.
Seq	Sequence 2857 BP;	626 A; 922 C; 699 G; 610 T;
Query Match	1/30;	Score 14; DB 12; Length 2857;
Best Local Similarity	100.0%;	Prod No. 317e-02; Indels 0; Caps 0;
Matches	14; Conservative	O: Mismatches 0; Indels 0; Caps 0;
Dd	2116 TCAGGAAAGCGC 2131	
Cp	896 TTCACGAAGCCAG 883	
RESULT	405	
ID	ID T03730 standard: cDNA; 2866 BP.	
ACC	T03730;	
DT	26-MAR-1996 (first entry)	
ORF	Transmembrane protein 55.11 cDNA	
Summary	Tumour necrosis factor receptor type 5, TNFR5, p55TC; intracellular domain binding protein; human immunodeficiency virus;	
Accession	U05854	
Keywords	HIV; ss.	
Source	Kemo sapiens.	
Location/Qualifiers	cds 4..2706	
FT	/*tag= a	
FT	/OS31544.h;	
FT	11-MAY-1995; U05854	
FT	11-MAY-1995; U05854	
FT	11-MAY-1994; IL-110632.	
FT	02-OCT-1994; IL-111163.	
FT	(GENB) VEDA RESEAR DEV CO LTD.	
FT	Boldin M, Mett I, Varfolomeev E, Wallach D;	
FT	NPJ: 96-010910/01.	
FT	TPSPR R88616	
FT	TPSPR R88616	
FT	useful for medulating receptor function, e.g., for treating tumors or HIV-infected cells	
FT	Claim 12; Fig 1a; 96pp: English.	
FT	Abstract: The present invention relates to a novel cDNA clone from a human fetal liver library (cDNA), which encodes a novel transmembrane protein, termed CD95, also referred to as FAS, which has been shown to be involved in apoptosis. This protein can be used to generate monoclonal antibodies, which may be used to study the role of this protein in apoptosis. The cDNA can be used to transform host cells or to facilitate homologous recombination.	
FT	Sequence 2866 BP; 688 A; 924 G; 674 T;	
Query Match	1/30;	Score 14; DB 12; Length 2866;
Best Local Similarity	100.0%;	Prod No. 317e-02; Indels 0; Caps 0;
Matches	14; Conservative	O: Mismatches 0; Indels 0; Caps 0;

[illegible]



























Tue Nov 17 08:55:26 1998

14-MAY-1993; US-062443  
 (UTV:) UNIV JEFFERSON THOMAS.  
 Canani E., Croce C;  
 WPI.: 95-006819/01.  
 NBT: R645666666 leukemia gene prods. - used for the  
 diagnosis and treatment of leukemias, partic. acute  
 lymphoblastic or nonlymphoblastic leukemia  
 Claim 8: Page 110-113; 30pp: English.  
 The invention relates to nucleotides derived from  
 the sequence of Q7468 and able to identify chromosome  
 abnormalities within the Af-9 gene of chromosome 9 are claimed.  
 The protein encoded by Af-9 encloses a nuclear targeting sequence  
 MKGK. The Af-9 protein is serine-rich (20%) including an  
 average of 12 phosphorylation sites, 42 acidic residues and has a higher than  
 average proline content.  
 Sequence 3176 BP: 1083 A: 713 G: 698 T:

```

Query Match      1.3%  Score 15;  DB 13;  Length 3376;
Best Local Similarity 100.0%;  Pred. No. 6.37e+01;
Matches 15;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

pb      312  ggaacgcagtaacat 326
          |||||
cp      83  ggaacgcagtaacat  69

```

RESULT	432
ID	ID
standard:	cDNA; 3415 BP.
CAC	736481;
DB-OCT-1996	(first entry) protein, beta-5, cDNA,
NCBI	HUMAN; integrin beta subunit; beta:5; carcinoma, lymphoid cell;
UniProt	Immunosassay; detection: mRNA; assay: ss.
SW	Homo sapiens.
Key	Location/Qualifiers
key	337-376
CDs	/tag= a
FI	signal_peptide 337..405
F1	/tag= b
F1	337..376
F1	/tag= c
FN	OS5527679-A.
PD	18-JUN-1996.
PR	18-MAY-1994.
PS	01-MAY-1994.
PPR	27-APR-1993.
PPR	US-054077.
PI	(DAND ) DANA FABER CANCER INST INC.
P1	Hemler ME, gelskov H;
P2	Nature 366, 685-690.
P-FSDB	P-FSDB 802134.

Tue Nov 17 08:55:26 1998  
US-08-887-977-9.mq

Db 1439 tgaagatcttgag 1452  
 |||||  
 Dy 956 TGAGATCTTGAG 969

RESULT 434  
 ID 780089 standard: cDNA: 3423 BP.

[illegible]

```

Query Match      1.3% Score 14: DB 34; Length 3423;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db          2568 gtagaagaagacta 2581
            |||.....|||
            |||.....|||
Py          982 CTCAGAGGAGACTA 985

RESULT 435
ID ID Q57072 standard; DNA; 3425 BP.
AC Q57072;
DE 26-JUL-1994 (first entry)
DEF AGC-modified DNA INS-8
DEF AGC-modification end products; AGC plasmids; transposon; ss.
KEY NUC MANUSCRIPT

```

Immunosay and RNA hybridisation assay for beta-5 protein - useful  
for the detection of carcinoma(s) and to distinguish different cell  
types  
Clam 11; Columns 13-20; 2lpp; English.  
The present sequence encodes the human integrin beta subunit  
CC  
CC found in carcinoma but not in lymphoid  
CC cells. An immunosay for the detection of carcinoma(s) can be  
CC contacting a sample with a monoclonal antibody (Ab) which binds a  
CC epitope of beta-5, and then with a labelled Ab (Ab) which binds 1  
CC epitope of beta-5 and detecting any bound label. An assay for  
CC hybridising to beta-5 mRNA and determining if it is capable of  
CC hybridising to the beta-5 cDNA and determining if it is capable of  
CC occurring. These assays are useful for detecting carcinoma(s), and for  
CC distinguishing between different cell types.  
Query Match 3415 BP: 752 A: 943 C: 1009 G: 711 T: 1.38; Score 14; DB 32; Length 3415

```
db      1869 cctgtccgcggagg 1882
          |||||
          |||||
dy      1013 CCTGTCCCGGAGG 1026
```

RESULT 433

ID ID7547 standard; DNA: 3420 BP.

DT 07-OCT-1997 (first entry)

DE Mortierella alpina cytochrome b5 genomic DNA.

CC Cytochrome b5; cytB5; preparation: production; human;

AB fatty acid; ss

OR Mortierella alpina.

PF J09121873-A.

PR 13-MAY-1997.

PS 203735.

SI AUG-1996; 203735.

FA (SUNR) SUPPLY LTD.

WI: 97-314231/29.

RT Mortierella alpina cytochrome b5 (cytB5) gene - used for recombinant production of cytb5 which is used in the preparation of essential Claim 9; Pages 9-10; 14pp; Japanese.

CS The present sequence encodes the Mortierella alpina cytochrome b5 (cytB5), useful for the efficient preparation of human essential fatty acids.

NC Sequence: 3420 BP; 820 A; 852 C; 858 G; 864 T.

Query Match 1.33; Score 14; DB 32; Length 3420;

Local Similarity 100.0%; Pred. No. 3.47e+02;

Tue Nov 17 08:55:26 1998  
119-08-887-977-9 .m

00340202599-A. 975 A. 787 C. 700 G. 981 T.  
18-JUL-1994: 3425 B3.  
19-JUL-1994: 006754.  
20-JUL-1994: 006754.  
22-JUL-1992: US-920385.  
(UTRQ) UNIV ROTENFELLER.  
Bucala RJ, Cerami A, Lee AT.  
1994. Advanced glycosylation end-products, typically in the form of  
age-pigments - can be transferred into cells and used to capture  
or activate transposons, e.g. to treat tumour cells  
Thompson R, Pig 4: 5pp. English.  
The expression and related cellular activity of a transposon TUS-9 that  
reacts with advanced glycosylation end products and is typically in  
the form of an AGE pigment that can be transferred into cells. The  
transposon may activate the transposons which  
are captured. Such capture and movement of transposons in a cell may  
be used to treat tumours.  
See also Q37055-73.  
Sequence 3425 B3.

```

Query Match      1.3% Score 14: DB 9; Length 3425;
Best Local Similarity 100.0%; Pred. No. 3.47e+02;
Matches          14; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Db              2637 TCACCAAAATAC 2640
|||||
DB              2637 TCACCAAAAATAC 2640

RESULT 436
ID 796832 standard; DNA; 3507 BP.
AC 796832;
CT 27-APR-1998 (first entry) tumour suppressor gene;
NCBI human pRB2/p130 gene;
NM Retinoblastoma protein;
KW cell cycle; tumour suppressor gene; cancer; molecular marker;
DIAGNOSIS: prognosis; predisposition; Endometrial carcinoma;
DISEASE: ovarian cancer; lung squamous cell carcinoma; lung adenocarcinoma;
SW ovary cancer; lung squamous cell carcinoma; lung adenocarcinoma;
WS Human; intron 2; db.
NS Human; intron 2; db.
NC P03738125-A1.
PD 16-OCT-1997.   00S598.
PF 03-APR-1997;   US-019512.
PG 03-APR-1997;   US-019512.
PH 05-JUN-1996;   US-019372.
PI 05-JUN-1996;   US-019372.
PJ 21-JUN-1996;   US-020196.
PK (UTJC.) UNIV JEFFERSON THOMAS.
PL Paid A. Lombardo A.
PM 8 JUN 1997.

```















Db 2113 ataatacttacc 2126  
Cp 27 ANTAACTTCACG 14

RESULT 450  
ID T60045 standard: DNA: 3738 BP.  
AC T60045:  
DT 14-MAY-1997 (first entry)  
DE 86Q3a toxin coding sequence  
KW Bacillus thuringiensis, hymenopteran pest, pharaoh ant;  
KW biological control, Monomorium pharaonis; delta-endotoxin, lepidoptera;  
KW insect; da.  
OS Bacillus thuringiensis isolate PS86Q3.  
PN 25359672 A  
PR 22-MAY-1991: US-703977.  
PR 22-MAY-1991: US-703977.  
PR 25-NOV-1991: US-797645.  
PR 24-NOV-1991: US-156232.  
PR 24-NOV-1991: US-156232.  
PA (MCO) MYCOGEN CORP.  
PI Foncegrada L, Fu J, Kennedy MK, Meier H, Payne JM,  
PI Randall JB, Schepf HE, Schwab GE, Ulick RU;  
PI P-PSDB: w18872.  
DT Bacillus thuringiensis toxin - active against hymenopteran pests  
DT Diclosure: Column 63-68; 64pp; English.  
CC This sequence represents the coding sequence for the 86Q3a toxin isolated  
CC from a strain of *Bacillus thuringiensis* isolated from a pharaoh ant colony.  
CC B.T. is a gram-positive spore forming, soil bacterium, characterized by  
CC parasporal crystalline protein inclusions. These proteins can be highly  
CC toxic to pests, and have been used to produce insect resistant plants.  
CC The previously isolated B.T. delta-endotoxins were mainly active against  
CC lepidopteran insects. This sequence encodes a protein that is active  
CC against hymenopteran insects. The protein encoded by this sequence is an  
CC example of a toxin of the invention, for which the sequences shown in  
CC w13888 and w13871 represent the generic formulae. As the toxins of the  
CC invention are active against hymenopteran pests, they can be used for the  
CC biological control of ants, particularly pharaoh ants (*Monomorium*  
CC pharaonis).

Query Match 1.38; Score 14; DB 13; Length 3760;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
P-PSDB: w18872

Db 2113 ataatacttacc 2126  
Cp 27 ANTAACTTCACG 14

RESULT 451  
ID Q31412 standard: DNA: 3738 BP.  
AC Q31412:  
DT 20-APR-1993 (first entry)  
DE Bacillus thuringiensis gene 86Q3(a).  
KW Bacillus thuringiensis isolate PS86Q3.  
KW Location/Qualifiers  
FH Key 1..3738  
FT cds  
FT 1..3738  
FT /note= "protein 86Q3(a)".  
PR 26-NOV-1992. A.  
PR 22-MAY-1992: US-703977.  
PR 22-MAY-1992: US-703977.  
PR 25-NOV-1991: US-797645.  
PR 12-MAY-1992: EP-104228.  
PA (MCO) MYCOGEN CORP.  
PI Kennedy MK, Meier H, Payne JM, Randall JB, Ulick RU;  
PI P-PSDB: w18872.  
DT Toxin proteins isolated from Bacillus thuringiensis - for controlling  
DT ants. e.g. fire, carpenter, Argentine and pharaoh ants  
DT Claim 15, Page 48; 71pp; English.  
CC This sequence represents the coding sequence for the toxin protein  
CC 86Q3(a). This is useful as a method of biological control of ants,  
CC e.g. fire ants, carpenter ants, Argentine ants and pharaoh ants,  
CC as an alternative to chemical insecticides.  
SQ Sequence 3738 BP; 1305 A; 602 C; 686 G; 1145 T;  
Query Match 1.38; Score 14; DB 5; Length 3738;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 2113 ataatacttacc 2126  
Cp 27 ANTAACTTCACG 14

RESULT 452  
ID Q75334 standard: cDNA: 3760 BP.  
AC Q75334:  
DT 24-NOV-1995 (first entry)  
DE Murine tyrosine kinase receptor, tie-2 cDNA.  
KW Murine tyrosine kinase receptor; tie-2; hematosis diagnosis; da.  
OS Mus musculus.  
FH Key 152..3523  
FT cds  
FT 152..3523  
FT /tag= a

FT signal\_peptide 152..211  
FT mat\_peptide 212..3520  
FT 704315382-A  
PR 15-NOV-1994.  
PR 06-MAY-1993: 129912.  
PR (SUDA) SUDA  
PR (SUDA) SUDA  
PR WPI: 95-03231/03.  
DR P-PSDB: R67391.  
DT DNA coding a tie-2 receptor and a tie-2 receptor - used in the  
DT diagnosing hematosis. 17pp; Japanese.  
CC This sequence encodes the murine tyrosine kinase receptor tie-2.  
CC Q75334 encodes R67391 the murine tyrosine kinase receptor, and can therefore  
CC be used in the diagnosis of hematosis.  
SQ Sequence 3760 BP; 1024 A; 876 C; 970 G; 890 T;  
Query Match 1.38; Score 14; DB 13; Length 3760;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 2833 agcattattacc 2846  
Cp 108 AGCATTATTACC 121

RESULT 453  
ID T61864 standard: DNA: 3780 BP.  
AC T61864:  
DT 06-JUN-1997 (first entry)  
DE c-Abl gene; antisense RNA; cancer; therapy; DNA damage;  
KW radiotherapy; mitomycin C; cytostatic; tyrosine kinase; sa.  
OS Homo sapiens.  
PN 90970818 A.  
PR 30-NOV-1995: US-520923.  
PR 30-NOV-1995: US-520923.  
PA (ARCH-) ARCH DEV CORP.  
PI (ARCH-) ARCH DEV CORP.  
PI WPI: 97-179167/16. EN, Weinrebbaum RR;  
DR Antisense RNA molecule that inhibits expression of c-Abl gene  
DR product - useful in killing malignant cells in cancer, optionally in  
DR combination with DNA damaging agent.  
CC The c-Abl gene (761864 and T61865) encodes a non-receptor tyrosine  
CC kinase that is activated DNA damaging agents. Antisense constructs

CC directed to the c-abl gene can be used to selectively inhibit the  
CC expression c-Abl following exposure of cells to a DNA damaging  
CC agent such as ionising radiation or mitomycin C. When used in  
CC combination with DNA damaging agents, the antisense RNA provides a  
CC synergistic killing effect useful for killing malignant cells  
CC in cancer treatment.  
SQ Sequence 3780 BP; 853 A; 1136 C; 1166 G; 625 T;  
Query Match 1.38; Score 14; DB 29; Length 3780;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 2737 gaagctctgag 2750  
Cp 602 GAAAGCTCTGAG 615

RESULT 454  
ID T95244 standard: cDNA: 3804 BP.  
AC T95244:  
DT 21-MAY-1998 (first entry)  
DE Human hematopoietic stem and progenitor cell antigen AC133 cDNA.  
KW AC133; hematopoietic stem cell; progenitor cell; antigen;  
OS marker; human; bone marrow; gene therapy; da.  
FH Key 38..2635  
FT cds  
FT 38..2635  
FT /tag= a  
FT sig\_peptide 38..94  
FT mat\_peptide 55..2632  
FT /tag= c  
PR 25-APR-1997: US-69930.  
PR 23-APR-1997: US-842382.  
PR 26-APR-1996: US-639891.  
PA (AMCE-) AMCELL CORP.  
PI (AMCE-) AMCELL CORP.  
PI WPI: 97-543778/50.  
DR P-PSDB: W26769.  
DT Antibody specific for hematopoietic stem/progenitor cell antigen.  
DT AC133 - useful to purify such cells, e.g. for use as gene therapy  
DT reagent for bone marrow transplantation  
DT Claim 21, Page 42-48; 78pp; English  
CC This claimed nucleic acid molecule includes a coding region for  
CC a human hematopoietic progenitor cell antigen (see W26769),  
CC designated AC133 antigen. It was isolated from a WMR-Rb-1  
CC cell line, which is a human myeloid leukemia cell line. The primer  
CC based on isolated peptides of purified AC133. The sequence of AC133  
CC is highly tissue specific. It is detected on a subset of







CC These were cloned into pBlac, an E. coli/PT shuttle vector made up  
 CC of replication origins from pBlac and pUC19. This was used to  
 CC transform E. coli DH5 $\alpha$ , and this grown on IPTG and XNA media.  
 CC White colonies were selected, purified, and either pMTG1627 contg.  
 CC the 2.7kb P817a toxin fragment or pMTG1628 contg. the 4.5kb P817b  
 CC toxin fragment (this sequence) isolated and sequenced. 1189 T;  
 CC Sequence 3867 BP; 1294 A; 635 C; 745 G; 1189 T;

Query Match 1.38; Score 14; DB 6; Length 3867;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttcaactgcgg 415  
 |||||||  
 QY 355 AACCTTACTGCG 368

## RESULT 459

ID Q30936 standard; DNA: 3867 BP.

AC Q30936;

DT 30-MAR-1993 (first entry)

KW nematode worms; nematocidal toxin; agriculture; plants;

OS crops; pests; Cry proteins.

NC Bacillus thuringiensis

FT Key Location/Qualifiers

CS 521 525

FT W09219739-A

PD 12-NOV-1992

PR 03-MAY-1991; US-630108

PR 31-JAN-1992; US-830050

PR 23-APR-1992; US-871510

PA (MCO) MYCOGEN CORP.

DR WPT: 92-39886/48

DR P-PSDB: R28810

DR New genes and toxins against nematodes - obt'd. from *Bacillus*

DR *thuringiensis* isolates with nematocidal activity

DR This sequence encodes the *Bacillus thuringiensis* delta-endotoxin

CC having nematocidal activity. A library was constructed in

CC lambdaDE3 using a strain P817 total cellular DNA partially

CC digested with *Sma*I and size fractionated by electrophoresis.

CC The library was screened with probe Q30949. Positive plaques were purified and

CC screened with probe Q30949. Positive plaques were purified and

CC resequenced, and positive plaques used to infect *E. coli* BW351 for

CC phase DNA preparation. Phase DNA was digested with *Eco*RI and

CC separated by electrophoresis and reprobbed with Q30949. Two positive

CC bands were present. Clone contg. 4.5 and 2.7 kb *B*st fragments.

CC

CC

CC

CC

CC These were cloned into pBlac, an E. coli/PT shuttle vector made up  
 CC of replication origins from pBlac and pUC19. This was used to  
 CC transform E. coli DH5 $\alpha$ , and this grown on IPTG and XNA media.  
 CC White colonies were selected, purified, and either pMTG1627 contg.  
 CC the 2.7kb P817a toxin fragment or pMTG1628 contg. the 4.5kb P817b  
 CC toxin fragment (this sequence) isolated and sequenced. 1189 T;  
 CC Sequence 3867 BP; 1294 A; 635 C; 745 G; 1189 T;

Query Match 1.38; Score 14; DB 5; Length 3867;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttcaactgcgg 415  
 |||||||  
 QY 355 AACCTTACTGCG 368

## RESULT 460

ID Q20335 standard; DNA: 3867 BP.

AC Q20335;

DT 26-MAR-1992 (first entry)

KW *B. thuringiensis* toxin gene P817b

OS nematocidal toxin; parasiticide; anthelmintic; parasite; ss.

NC *Bacillus thuringiensis*.

FT Key Location/Qualifiers

CS 521 525

FT W09219739-A

PD 12-NOV-1992

PR 03-MAY-1991; US-630108

PR 31-JAN-1992; US-830050

PR 23-APR-1992; US-871510

PA (MCO) MYCOGEN CORP.

DR WPT: 92-39886/48

DR P-PSDB: R28810

DR New genes and toxins against nematodes - obt'd. from *Bacillus*

DR *thuringiensis* isolates with nematocidal activity

DR This sequence encodes the *Bacillus thuringiensis* delta-endotoxin

CC having nematocidal activity. A library was constructed in

CC lambdaDE3 using a strain P817 total cellular DNA partially

CC digested with *Sma*I and size fractionated by electrophoresis.

CC The library was screened with probe Q30949. Positive plaques were purified and

CC screened with probe Q30949. Positive plaques were purified and

CC resequenced, and positive plaques used to infect *E. coli* BW351 for

CC phase DNA preparation. Phase DNA was digested with *Eco*RI and

CC separated by electrophoresis and reprobbed with Q30949. Two positive

CC bands were present. Clone contg. 4.5 and 2.7 kb *B*st fragments.

CC

CC

CC

CC

Query Match 1.38; Score 14; DB 3; Length 3867;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttcaactgcgg 415  
 |||||||  
 QY 355 AACCTTACTGCG 368

## RESULT 461

ID Q94054 standard; DNA: 3867 BP.

AC Q94054;

DT 16-MAY-1995 (first entry)

KW Insecticide; toxin; acaricide-active toxin; delta-endotoxin; P817b; P817;

OS two spotted spider mite; *Tetranychus urticae*; *Bacillus thuringiensis*; da.

NC *Bacillus thuringiensis* strain P817.

FT Key Location/Qualifiers

CS 521 525

FT W09219739-A

PD 12-NOV-1992

PR 03-MAY-1991; US-630108

PR 31-JAN-1992; US-830050

PR 23-APR-1992; US-871510

PA (MCO) MYCOGEN CORP.

DR WPT: 92-39886/48

DR P-PSDB: R28810

DR New genes and toxins against nematodes - obt'd. from *Bacillus*

DR *thuringiensis* isolates with nematocidal activity

DR This sequence encodes the *Bacillus thuringiensis* delta-endotoxin

CC having nematocidal activity. A library was constructed in

CC lambdaDE3 using a strain P817 total cellular DNA partially

CC digested with *Sma*I and size fractionated by electrophoresis.

CC The library was screened with probe Q30949. Positive plaques were purified and

CC screened with probe Q30949. Positive plaques were purified and

CC resequenced, and positive plaques used to infect *E. coli* BW351 for

CC phase DNA preparation. Phase DNA was digested with *Eco*RI and

CC separated by electrophoresis and reprobbed with Q30949. Two positive

CC bands were present. Clone contg. 4.5 and 2.7 kb *B*st fragments.

CC

CC

CC

CC

CC

CC

CC

Query Match 1.38; Score 14; DB 16; Length 3867;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttcaactgcgg 415  
 |||||||  
 QY 355 AACCTTACTGCG 368

## RESULT 462

ID Q71180 standard; DNA: 3867 BP.

AC Q71180;

DT 16-MAY-1995 (first entry)

KW *Bacillus thuringiensis* P817b delta-endotoxin gene.

OS Endotoxin; *Bacillus thuringiensis*; P817; toxin; delta-endotoxin;

NC parasiticide; crystalline; inclusion; toxic; pest; spider mite;

FT Key Location/Qualifiers

CS 521 525

FT W09219739-A

PD 12-NOV-1992

PR 03-MAY-1991; US-630108

PR 31-JAN-1992; US-830050

PR 23-APR-1992; US-871510

PA (MCO) MYCOGEN CORP.

DR WPT: 92-39886/48

DR P-PSDB: R28810

DR New genes and toxins against nematodes - obt'd. from *Bacillus*

DR *thuringiensis* isolates with nematocidal activity

DR This sequence encodes the *Bacillus thuringiensis* delta-endotoxin

CC having nematocidal activity. A library was constructed in

CC lambdaDE3 using a strain P817 total cellular DNA partially

CC digested with *Sma*I and size fractionated by electrophoresis.

CC The library was screened with probe Q30949. Positive plaques were purified and

CC screened with probe Q30949. Positive plaques were purified and

CC resequenced, and positive plaques used to infect *E. coli* BW351 for

CC phase DNA preparation. Phase DNA was digested with *Eco*RI and

CC separated by electrophoresis and reprobbed with Q30949. Two positive

CC bands were present. Clone contg. 4.5 and 2.7 kb *B*st fragments.

CC

CC

CC

CC

CC

CC

CC

Query Match 1.38; Score 14; DB 3; Length 3867;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttcaactgcgg 415  
 |||||||  
 QY 355 AACCTTACTGCG 368

## RESULT 461

ID Q94054 standard; DNA: 3867 BP.

AC Q94054;

DT 16-MAY-1995 (first entry)

KW Insecticide; toxin; acaricide-active toxin; delta-endotoxin; P817b; P817;

OS two spotted spider mite; *Tetranychus urticae*; *Bacillus thuringiensis*; da.

NC *Bacillus thuringiensis* strain P817.

FT Key Location/Qualifiers

CS 521 525

FT W09219739-A

PD 12-NOV-1992

PR 03-MAY-1991; US-630108

PR 31-JAN-1992; US-830050

PR 23-APR-1992; US-871510

PA (MCO) MYCOGEN CORP.

DR WPT: 92-39886/48

DR P-PSDB: R28810

DR New genes and toxins against nematodes - obt'd. from *Bacillus*

DR *thuringiensis* isolates with nematocidal activity

DR This sequence encodes the *Bacillus thuringiensis* delta-endotoxin

CC having nematocidal activity. A library was constructed in

CC lambdaDE3 using a strain P817 total cellular DNA partially

CC digested with *Sma*I and size fractionated by electrophoresis.

CC The library was screened with probe Q30949. Positive plaques were purified and

CC screened with probe Q30949. Positive plaques were purified and

CC resequenced, and positive plaques used to infect *E. coli* BW351 for

CC phase DNA preparation. Phase DNA was digested with *Eco*RI and

CC separated by electrophoresis and reprobbed with Q30949. Two positive

CC bands were present. Clone contg. 4.5 and 2.7 kb *B*st fragments.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC



DB	402	acccttaactggg	415		
AC	Q31410;	standard; DNA;	3867 BP.		
TC	Q31410;				
DT	20-APR-1993	(first entry)			
DE	Bacillus thuringiensis gene 17b.				
OS	Bacterial protein toxin delta strain ps17.				
OG	Bacillus thuringiensis				
FH	Key Location/Qualifiers				
FT	cds	1..3867			
PT	/tag= "a"				
PF	/note= "protein 17b"				
PP	W09220802.A.				
PD	26-NOV-1992.				
PE	22-MAY-1992.	004316.			
PR	25-NOV-1991.	US-792637.			
PR	25-NOV-1991.	US-792645.			
PR	12-MAY-1992.	EP-304228.			
PA	(MTCO) MTCOGEN CORP.				
P1	Kennedy MT, Meier H, Payne JM, Randall JB, Dick BW:				
P2	Protein toxins from Bacillus thuringiensis - for controlling				
DR	p-PDMS: R32097/30.				
DT	Toxin proteins isolated from Bacillus thuringiensis - for controlling				
PT	ants. e.g. fire, carpenter, argentine and pharaoh ants				
PT	Disclosure, Page 39; 7:pp: English.				
CC	This is useful as a method of biological control of ants, e.g. fire				
CC	ants, carpenter ants, argentine ants and pharaoh ants, as an				
CC	alternative to chemical insecticides.				
CC	Sequence 3867 BP: 1397 A; 635 C; 751 G; 1189 T.				
Query Match	1.3%; Score 14; DB 5; Length 3867;				
Best Local Similarity	100.0%; Pred.No. 3.47e+02;				
Matches	14; Conservative	O; Mismatches	O; Indels	O; Gaps	O;
DB	402	acccttaactggg	415		
AC	Q31410;	standard; DNA;	3867 BP.		
TC	Q31410;				
DT	20-APR-1993	(first entry)			
DE	Bacillus thuringiensis gene 17b.				
OS	Bacterial protein toxin delta strain ps17.				
OG	Bacillus thuringiensis				
FH	Key Location/Qualifiers				
FT	cds	1..3867			
PT	/tag= "a"				
PF	/note= "protein 17b"				
PP	W09220802.A.				
PD	26-NOV-1992.				
PE	22-MAY-1992.	004316.			
PR	25-NOV-1991.	US-792637.			
PR	25-NOV-1991.	US-792645.			
PR	12-MAY-1992.	EP-304228.			
PA	(MTCO) MTCOGEN CORP.				
P1	Kennedy MT, Meier H, Payne JM, Randall JB, Dick BW:				
P2	Protein toxins from Bacillus thuringiensis - for controlling				
DR	p-PDMS: R32097/30.				
DT	Toxin proteins isolated from Bacillus thuringiensis - for controlling				
PT	ants. e.g. fire, carpenter, argentine and pharaoh ants				
PT	Disclosure, Page 39; 7:pp: English.				
CC	This is useful as a method of biological control of ants, e.g. fire				
CC	ants, carpenter ants, argentine ants and pharaoh ants, as an				
CC	alternative to chemical insecticides.				
CC	Sequence 3867 BP: 1397 A; 635 C; 751 G; 1189 T.				
Query Match	1.3%; Score 14; DB 5; Length 3867;				
Best Local Similarity	100.0%; Pred.No. 3.47e+02;				
Matches	14; Conservative	O; Mismatches	O; Indels	O; Gaps	O;
DB	402	acccttaactggg	415		
AC	Q31410;	standard; DNA;	3867 BP.		
TC	Q31410;				
DT	20-APR-1993	(first entry)			
DE	Bacillus thuringiensis gene 17b.				
OS	Bacterial protein toxin delta strain ps17.				
OG	Bacillus thuringiensis				
FH	Key Location/Qualifiers				
FT	cds	1..3867			
PT	/tag= "a"				
PF	/note= "protein 17b"				
PP	W09220802.A.				
PD	26-NOV-1992.				
PE	22-MAY-1992.	004316.			
PR	25-NOV-1991.	US-792637.			
PR	25-NOV-1991.	US-792645.			
PR	12-MAY-1992.	EP-304228.			
PA	(MTCO) MTCOGEN CORP.				
P1	Kennedy MT, Meier H, Payne JM, Randall JB, Dick BW:				
P2	Protein toxins from Bacillus thuringiensis - for controlling				
DR	p-PDMS: R32097/30.				
DT	Toxin proteins isolated from Bacillus thuringiensis - for controlling				
PT	ants. e.g. fire, carpenter, argentine and pharaoh ants				
PT	Disclosure, Page 39; 7:pp: English.				
CC	This is useful as a method of biological control of ants, e.g. fire				
CC	ants, carpenter ants, argentine ants and pharaoh ants, as an				
CC	alternative to chemical insecticides.				
CC	Sequence 3867 BP: 1397 A; 635 C; 751 G; 1189 T.				
Query Match	1.3%; Score 14; DB 5; Length 3867;				
Best Local Similarity	100.0%; Pred.No. 3.47e+02;				
Matches	14; Conservative	O; Mismatches	O; Indels	O; Gaps	O;
DB	402	acccttaactggg	415		
AC	Q31410;	standard; DNA;	3867 BP.		
TC	Q31410;				
DT	20-APR-1993	(first entry)			
DE	Bacillus thuringiensis gene 17b.				
OS	Bacterial protein toxin delta strain ps17.				
OG	Bacillus thuringiensis				
FH	Key Location/Qualifiers				
FT	cds	1..3867			
PT	/tag= "a"				
PF	/note= "protein 17b"				
PP	W09220802.A.				
PD	26-NOV-1992.				
PE	22-MAY-1992.	004316.			
PR	25-NOV-1991.	US-792637.			
PR	25-NOV-1991.	US-792645.			
PR	12-MAY-1992.	EP-304228.			

```

FF 14-MAR-1990; 063306.  
PA CDL1041  
PA (SHIS ) SHISEIDO KK.  
WPI: 92-012701/02.  
P-PDSB: R32028.  
DR OUTD. Outdones coding for peptide C terminal amidation enzyme  
DR P-PDSB: R32028.  
PS Claim 1: Fig 1: 18pp; Japanese.  
CC The sequence was determined from a clone isolated from a library  
CC prep'd. from RNA extracted from horse atrium tissue. The sequence  
CC encodes a protein of 394 aa. The sequence is identical to the  
CC sequence up to base 2437 where they diverge, encoding different  
CC C-terminal. The different termini are created by deletions in the  
CC last portion of DNA (see feature table). The sequence encodes an  
CC monomer (see seq. tab. 02019) which has been shown to have  
CC sequence 3904 aa; 1137 A; 752 C; 888 G; 1088 T;  
Query Match 1.3%; Score 15; DB 3; Length 3904;  
Mat.Local Similarity 100.0%; Pred.No. 8.3%e=0.1; Indels 0; Gaps 0;  
Date 15/11/91; Conservative 0; Mismatches 0.  
AB 3687 tgccttcacaaaaca 3701  
QY 522 TGCTTCGACCAAAA 536  
  
RESULT 467  
CD Q11041 standard; DNA; 3925 BP.  
AK Q11041  
DT 17-MAR-1991 (first entry)  
DD Encodes equine C-terminal  
RW C-terminal amidating enzyme: calcitonin; gastrin; ss.  
RS Swiss catubulin. Location/Qualifiers  
FH Key  
FT cds 11..2953  
FT /*tag = s  
FT /Product=C-terminal amidating enzyme  
W9102790-A  
PD 07-MAR-1991.  
PP 14-AUG-1990; J01036.  
PD 15-AUG-1989; JP-285697.  
PR 15-AUG-1989; JP-285697.  
PR 26-MAR-1990; JP-076331.  
PR 24-APR-1990; JP-106412.  
PR 02-AUG-1990; JP-205475.  
PR 02-AUG-1990; JP-205475.  
PS (SHIS ) SHISEIDO KK.  
WPI: 91-08774/12.  
P-PDSB: R11113.  
DR DR New enzymes useful in peptide C-amidation - convert peptide C-  
PT terminal glycine to alpha hydroxyglycine then remove this leaving  
PT

```



PF Disclosure: Fig 13: 11pp; Japanese.  
 CC cloned using probes derived from sequences encoding the rat enzyme.  
 CC The C-terminal amidating enzyme can be used to amidate  
 CC physiologically active peptides such as calcitonin and gastrin.  
 CC It also amidates 0-Tyr-Tyr-Gly, Phe-Gly-Phe-Gly and Gly-Phe-Gly.  
 CC Sequence 3225 BP; 1146 A; 789 C; 899 G; 1091 T;  
 SQ

Query Match  
 Best Local Similarity 100.04; Pred. No. 8.37e+01; Length 3925;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3708 tgcctccaccacaa 3722  
 QY 522 tgcctccaccacaa 536

## RESULT 468

ID 001736 standard; cDNA: 3925 BP.  
 AC 16-AUG-1990 (first entry)  
 DT 16-AUG-1990 (first entry)  
 DE Sequence of plasmid pcR11 encoding cDNA of fowl prolactin.  
 DR Key  
 PR Prolactin: prolactin; pituitary; fowl; poultry; da.  
 PA (NIGP-) Nigpon Gene KK.  
 P-PSB8-203913.  
 DR P-PSB8-203913.  
 PT Recombinant avian prolactin and recombinant avian prolactin -  
 PT used for avian breeding and proliferation.  
 PS Disclosure: Fig 1A-G; Pp: Japanese.  
 CC The mature prolactin derived was used to induce maternal behaviour in  
 CC the mouse.  
 CC Fowl and encourage breeding and proliferation.  
 CC Sequence 3925 BP; 1047 A; 966 C; 938 G; 975 T;  
 SQ

Query Match  
 Best Local Similarity 100.04; Pred. No. 8.37e+01;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 469

ID 001736 standard; cDNA: 3925 BP.  
 AC 16-AUG-1990 (first entry)  
 DT 16-AUG-1990 (first entry)  
 DE Sequence of plasmid pcR11 encoding cDNA of fowl prolactin.  
 DR Key  
 PR Prolactin: prolactin; pituitary; fowl; poultry; da.  
 PA (NIGP-) Nigpon Gene KK.  
 P-PSB8-203913.  
 DR P-PSB8-203913.  
 PT Recombinant avian prolactin and recombinant avian prolactin -  
 PT used for avian breeding and proliferation.  
 PS Disclosure: Fig 1A-G; Pp: Japanese.  
 CC The mature prolactin derived was used to induce maternal behaviour in  
 CC the mouse.  
 CC Fowl and encourage breeding and proliferation.  
 CC Sequence 3925 BP; 1047 A; 966 C; 938 G; 975 T;  
 SQ

OS Homo sapiens.  
 PF Key  
 PH Key  
 FT cds  
 FT 304..3531  
 FT 1756..1758  
 FT /tag= b  
 FT /note= "bases 1756-1758 may be replaced by gta,  
 FT 98a or 99a"  
 FT /tag= c  
 FT /note= "bases 3090-3092 may be replaced by agt,  
 FT 98a or 99a"

FT W0502053-A  
 FT 05-JUL-1994: US-088483.  
 FT 06-JUL-1994: US-088483.  
 FT 30-JUL-1994: US-099998.  
 FT 30-JUL-1994: US-113971.  
 FT 30-AUG-1994: US-113971.  
 FT 04-APR-1994: US-223626.  
 FT (SCHE) SCHERING CORP.  
 FT Arai N, Masuda ES, Tokumitsu H;  
 FT WPI:US-06699/09.  
 DT Nuclear factor of Activated T cells 120 protein, antibody and  
 DT nuclear acid - useful as transcriptional regulator of genes, such  
 DT as cytokines

PF Disclosure: Page 76-93; 10pp; English.  
 CC given in 084300 and encoded NF-AT120 (Nuclear Factor of  
 CC Activated T cells protein 120) (R66873). Amplified DNA was used  
 CC to screen for related NF-AT proteins by hybridization, leading to  
 CC the discovery of 3 related subfamilies, designated class C, P and  
 CC X (R66874, R66875, R66876, R66877, R66878, R66879, R66880, R66881, R66882, R66883, R66884, R66885, R66886, R66887, R66888, R66889, R66890, R66891, R66892, R66893, R66894, R66895, R66896, R66897, R66898, R66899, R66900, R66901, R66902, R66903, R66904, R66905, R66906, R66907, R66908, R66909, R66910, R66911, R66912, R66913, R66914, R66915, R66916, R66917, R66918, R66919, R66920, R66921, R66922, R66923, R66924, R66925, R66926, R66927, R66928, R66929, R66930, R66931, R66932, R66933, R66934, R66935, R66936, R66937, R66938, R66939, R66940, R66941, R66942, R66943, R66944, R66945, R66946, R66947, R66948, R66949, R66950, R66951, R66952, R66953, R66954, R66955, R66956, R66957, R66958, R66959, R66960, R66961, R66962, R66963, R66964, R66965, R66966, R66967, R66968, R66969, R66970, R66971, R66972, R66973, R66974, R66975, R66976, R66977, R66978, R66979, R66980, R66981, R66982, R66983, R66984, R66985, R66986, R66987, R66988, R66989, R66990, R66991, R66992, R66993, R66994, R66995, R66996, R66997, R66998, R66999, R67000, R67001, R67002, R67003, R67004, R67005, R67006, R67007, R67008, R67009, R67010, R67011, R67012, R67013, R67014, R67015, R67016, R67017, R67018, R67019, R67020, R67021, R67022, R67023, R67024, R67025, R67026, R67027, R67028, R67029, R67030, R67031, R67032, R67033, R67034, R67035, R67036, R67037, R67038, R67039, R67040, R67041, R67042, R67043, R67044, R67045, R67046, R67047, R67048, R67049, R67050, R67051, R67052, R67053, R67054, R67055, R67056, R67057, R67058, R67059, R67060, R67061, R67062, R67063, R67064, R67065, R67066, R67067, R67068, R67069, R67070, R67071, R67072, R67073, R67074, R67075, R67076, R67077, R67078, R67079, R67080, R67081, R67082, R67083, R67084, R67085, R67086, R67087, R67088, R67089, R67090, R67091, R67092, R67093, R67094, R67095, R67096, R67097, R67098, R67099, R67100, R67101, R67102, R67103, R67104, R67105, R67106, R67107, R67108, R67109, R67110, R67111, R67112, R67113, R67114, R67115, R67116, R67117, R67118, R67119, R67120, R67121, R67122, R67123, R67124, R67125, R67126, R67127, R67128, R67129, R67130, R67131, R67132, R67133, R67134, R67135, R67136, R67137, R67138, R67139, R67140, R67141, R67142, R67143, R67144, R67145, R67146, R67147, R67148, R67149, R67150, R67151, R67152, R67153, R67154, R67155, R67156, R67157, R67158, R67159, R67160, R67161, R67162, R67163, R67164, R67165, R67166, R67167, R67168, R67169, R67170, R67171, R67172, R67173, R67174, R67175, R67176, R67177, R67178, R67179, R67180, R67181, R67182, R67183, R67184, R67185, R67186, R67187, R67188, R67189, R67190, R67191, R67192, R67193, R67194, R67195, R67196, R67197, R67198, R67199, R67200, R67201, R67202, R67203, R67204, R67205, R67206, R67207, R67208, R67209, R67210, R67211, R67212, R67213, R67214, R67215, R67216, R67217, R67218, R67219, R67220, R67221, R67222, R67223, R67224, R67225, R67226, R67227, R67228, R67229, R67230, R67231, R67232, R67233, R67234, R67235, R67236, R67237, R67238, R67239, R67240, R67241, R67242, R67243, R67244, R67245, R67246, R67247, R67248, R67249, R67250, R67251, R67252, R67253, R67254, R67255, R67256, R67257, R67258, R67259, R67260, R67261, R67262, R67263, R67264, R67265, R67266, R67267, R67268, R67269, R67270, R67271, R67272, R67273, R67274, R67275, R67276, R67277, R67278, R67279, R67280, R67281, R67282, R67283, R67284, R67285, R67286, R67287, R67288, R67289, R67290, R67291, R67292, R67293, R67294, R67295, R67296, R67297, R67298, R67299, R67300, R67301, R67302, R67303, R67304, R67305, R67306, R67307, R67308, R67309, R67310, R67311, R67312, R67313, R67314, R67315, R67316, R67317, R67318, R67319, R67320, R67321, R67322, R67323, R67324, R67325, R67326, R67327, R67328, R67329, R67330, R67331, R67332, R67333, R67334, R67335, R67336, R67337, R67338, R67339, R67340, R67341, R67342, R67343, R67344, R67345, R67346, R67347, R67348, R67349, R67350, R67351, R67352, R67353, R67354, R67355, R67356, R67357, R67358, R67359, R67360, R67361, R67362, R67363, R67364, R67365, R67366, R67367, R67368, R67369, R67370, R67371, R67372, R67373, R67374, R67375, R67376, R67377, R67378, R67379, R67380, R67381, R67382, R67383, R67384, R67385, R67386, R67387, R67388, R67389, R67390, R67391, R67392, R67393, R67394, R67395, R67396, R67397, R67398, R67399, R67400, R67401, R67402, R67403, R67404, R67405, R67406, R67407, R67408, R67409, R67410, R67411, R67412, R67413, R67414, R67415, R67416, R67417, R67418, R67419, R67420, R67421, R67422, R67423, R67424, R67425, R67426, R67427, R67428, R67429, R67430, R67431, R67432, R67433, R67434, R67435, R67436, R67437, R67438, R67439, R67440, R67441, R67442, R67443, R67444, R67445, R67446, R67447, R67448, R67449, R67450, R67451, R67452, R67453, R67454, R67455, R67456, R67457, R67458, R67459, R67460, R67461, R67462, R67463, R67464, R67465, R67466, R67467, R67468, R67469, R67470, R67471, R67472, R67473, R67474, R67475, R67476, R67477, R67478, R67479, R67480, R67481, R67482, R67483, R67484, R67485, R67486, R67487, R67488, R67489, R67490, R67491, R67492, R67493, R67494, R67495, R67496, R67497, R67498, R67499, R67500, R67501, R67502, R67503, R67504, R67505, R67506, R67507, R67508, R67509, R67510, R67511, R67512, R67513, R67514, R67515, R67516, R67517, R67518, R67519, R67520, R67521, R67522, R67523, R67524, R67525, R67526, R67527, R67528, R67529, R67530, R67531, R67532, R67533, R67534, R67535, R67536, R67537, R67538, R67539, R67540, R67541, R67542, R67543, R67544, R67545, R67546, R67547, R67548, R67549, R67550, R67551, R67552, R67553, R67554, R67555, R67556, R67557, R67558, R67559, R67560, R67561, R67562, R67563, R67564, R67565, R67566, R67567, R67568, R67569, R67570, R67571, R67572, R67573, R67574, R67575, R67576, R67577, R67578, R67579, R67580, R67581, R67582, R67583, R67584, R67585, R67586, R67587, R67588, R67589, R67590, R67591, R67592, R67593, R67594, R67595, R67596, R67597, R67598, R67599, R67600, R67601, R67602, R67603, R67604, R67605, R67606, R67607, R67608, R67609, R67610, R67611, R67612, R67613, R67614, R67615, R67616, R67617, R67618, R67619, R67620, R67621, R67622, R67623, R67624, R67625, R67626, R67627, R67628, R67629, R67630, R67631, R67632, R67633, R67634, R67635, R67636, R67637, R67638, R67639, R67640, R67641, R67642, R67643, R67644, R67645, R67646, R67647, R67648, R67649, R67650, R67651, R67652, R67653, R67654, R67655, R67656, R67657, R67658, R67659, R67660, R67661, R67662, R67663, R67664, R67665, R67666, R67667, R67668, R67669, R67670, R67671, R67672, R67673, R67674, R67675, R67676, R67677, R67678, R67679, R67680, R67681, R67682, R67683, R67684, R67685, R67686, R67687, R67688, R67689, R67690, R67691, R67692, R67693, R67694, R67695, R67696, R67697, R67698, R67699, R67700, R67701, R67702, R67703, R67704, R67705, R67706, R67707, R67708, R67709, R67710, R67711, R67712, R67713, R67714, R67715, R67716, R67717, R67718, R67719, R67720, R67721, R67722, R67723, R67724, R67725, R67726, R67727, R67728, R67729, R67730, R67731, R67732, R67733, R67734, R67735, R67736, R67737, R67738, R67739, R67740, R67741, R67742, R67743, R67744, R67745, R67746, R67747, R67748, R67749, R67750, R67751, R67752, R67753, R67754, R67755, R67756, R67757, R67758, R67759, R67760, R67761, R67762, R67763, R67764, R67765, R67766, R67767, R67768, R67769, R67770, R67771, R67772, R67773, R67774, R67775, R67776, R67777, R67778, R67779, R67780, R67781, R67782, R67783, R67784, R67785, R67786, R67787, R67788, R67789, R67790, R67791, R67792, R67793, R67794, R67795, R67796, R67797, R67798, R67799, R67800, R67801, R67802, R67803, R67804, R67805, R67806, R67807, R67808, R67809, R67810, R67811, R67812, R67813, R67814, R67815, R67816, R67817, R67818, R67819, R67820, R67821, R67822, R67823, R67824, R67825, R67826, R67827, R67828, R67829, R67830, R67831, R67832, R67833, R67834, R67835, R67836, R67837, R67838, R67839, R67840, R67841, R67842, R67843, R67844, R67845, R67846, R67847, R67848, R67849, R67850, R67851, R67852, R67853, R67854, R67855, R67856, R67857, R67858, R67859, R67860, R67861, R67862, R67863, R67864, R67865, R67866, R67867, R67868, R67869, R67870, R67871, R67872, R67873, R67874, R67875, R67876, R67877, R67878, R67879, R67880, R67881, R67882, R67883, R67884, R67885, R67886, R67887, R67888, R67889, R67890, R67891, R67892, R67893, R67894, R67895, R67896, R67897, R67898, R67899, R67900, R67901, R67902, R67903, R67904, R67905, R67906, R67907, R67908, R67909, R67910, R67911, R67912, R67913, R67914, R67915, R67916, R67917, R67918, R67919, R67920, R67921, R67922, R67923, R67924, R67925, R67926, R67927, R67928, R67929, R67930, R67931, R67932, R67933, R67934, R67935, R67936, R67937, R67938, R67939, R67940, R67941, R67942, R67943, R67944, R67945, R67946, R67947, R67948, R67949, R67950, R67951, R67952, R67953, R67954, R67955, R67956, R67957, R67958, R67959, R67960, R67961, R67962, R67963, R67964, R67965, R67966, R67967, R67968, R67969, R67970, R67971, R67972, R67973, R67974, R67975, R67976, R67977, R67978, R67979, R67980, R67981, R67982, R67983, R67984, R67985, R67986, R67987, R67988, R67989, R67990, R67991, R67992, R67993, R67994, R67995, R67996, R67997, R67998, R67999, R68000, R68001, R68002, R68003, R68004, R68005, R68006, R68007, R68008, R68009, R68010, R68011, R68012, R68013, R68014, R68015, R68016, R68017, R68018, R68019, R68020, R68021, R68022, R68023, R68024, R68025, R68026, R68027, R68028, R68029, R68030, R68031, R68032, R68033, R68034, R68035, R68036, R68037, R68038, R68039, R68040, R68041, R68042, R68043, R68044, R68045, R68046, R68047, R68048, R68049, R68050, R68051, R68052, R68053, R68054, R68055, R68056, R68057, R68058, R68059, R68060, R68061, R68062, R68063, R68064, R68065, R68066, R68067, R68068, R68069, R68070, R68071, R68072, R68073, R68074, R68075, R68076, R68077, R68078, R68079, R68080, R68081, R68082, R68083, R68084, R68085, R68086, R68087, R68088, R68089, R68090, R68091, R68092, R68093, R68094, R68095, R68096, R68097, R68098, R68099, R68100, R68101, R68102, R68103, R68104, R68105, R68106, R68107, R68108, R68109, R68110, R68111, R68112, R68113, R68114, R68115, R68116, R68117, R68118, R68119, R68120, R68121, R68122, R68123, R68124, R68125, R68126, R68127, R68128, R68129, R68130, R68131, R68132, R68133, R68134, R68135, R68136, R68137, R68138, R68139, R68140, R68141, R68142, R68143, R68144, R68145, R68146, R68147, R68148, R68149, R68150, R68151, R68152, R68153, R68154, R68155, R68156, R68157, R68158, R68159, R68160, R68161, R68162, R68163, R68164, R68165, R68166, R68167, R68168, R68169, R68170, R68171, R68172, R68173, R68174, R68175, R68176, R68177, R68178, R68179, R68180, R68181, R68182, R68183, R68184, R68185, R68186, R68187, R68188, R68189, R68190, R68191, R68192, R68193, R68194, R68195, R68196, R68197, R68198, R68199, R68200, R68201, R68202, R68203, R68204, R68205, R68206, R68207, R68208, R68209, R68210, R68211, R68212, R68213, R68214, R68215, R68216, R68217, R68218, R68219, R68220, R68221, R68222, R68223, R68224, R68225, R68226, R68227, R68228, R68229, R68230, R68231, R68232, R68233, R68234, R68235, R68236, R68237, R68238, R68239, R68240, R68241, R68242, R68243, R68244, R68245, R68246, R68247, R68248, R68249, R68250, R68251, R68252, R68253, R68254, R68255, R68256, R68257, R68258, R68259, R68260, R68261, R68262, R68263, R68264, R68265, R68266, R68267, R68268, R68269, R68270, R68271, R68272, R68273, R68274, R68275, R68276, R68277, R68278, R68279, R68280, R68281, R68282, R68283, R68284, R68285, R68286, R68287, R68288, R68289, R68290, R68291, R68292, R68293, R68294, R68295, R68296, R68297, R68298, R68299, R68300, R68301, R68302, R68303, R68304, R68305, R68306, R68307, R68308, R68309, R68310, R68311, R68312, R68313, R68314, R68315, R68316, R68317, R68318, R68319, R68320, R68321, R68322, R68323, R68324, R68325, R68326, R68327, R68328, R68329, R68330, R68331, R68332, R68333, R68334, R68335, R68336, R68337, R68338, R68339, R68340, R68341, R68342, R68343, R68344, R68345, R68346, R68347, R68348, R68349, R68350, R68351, R6



115-08-887-077-0-000

58)•

68)

pendent terminator."

pendent terminator.



FT		/tag= 5	
PT		/note== deleted in Q20199 and Q20200*	
PN	J03162484-A.		
PD	22-NOV-1991.	051906	
PZ	14-MAR-1990.	JP-0653106.	
PA	(SHIS ) SHISEIDO KK.		
PP	WPI: 92-012701/02.		
DR	P-RSDS: R3202S.		
DT	obtd. from base coding for peptide C terminal amidation enzyme		
PT	Claim 1; Fig 1; 18pp; Japanese.		
CC	The sequence was determined from a clone isolated from a library		
CC	of cDNA fragments extracted from mouse atrium tissue. The sequence		
CC	is identical to that reported by Horiuchi et al. (1990) except for		
CC	sequence up to base 2417 where they diverge, encoding different		
CC	C-terminal. The different termini are created by deletions in the		
CC	last portion of DNA (see feature table). The sequence encodes an		
CC	enzyme having EC number 3.4.21.99 (see feature table).		
SQ	Sequence 4045 BP; 1181 Aa; 825 C; 922 G; 1117 T;		
	Query Match similarity 100%; Pos. No. 354-401; Caps 0;		
	Matches total 15; Conservative 0; Mismatches 0; Caps 0;		
DB	3828 tgccttcacacaaa 3942		
CT			
	532 tgccttcacacaaa 536		
RESULT	477		
AD	V10628		
AC	standard; CONDA: 4066 BP.		
DT	23-JUN-1998 (first entry)		
DE	Human glycogen phosphorylase CDNA.		
DD	Glycogen phosphorylase; human; brain; retinal pigment epithelium cells;		
KW	TRAP; disease: diagnosis; antibody: cytotoxic T cell ds		
FF	Homo sapiens.		
FS	Key	Location/Qualifiers	
CH	CDS	1..4066	
		/tag= 5	
FT		/product= brain glycogen phosphorylase	
PN	M09748947-AI.		
PD	25-JUN-1997		
PZ	25-JUN-1997	U011089	
PA	(LUDM-) LUDWIG INST CANCER RES.		
PP	REGC.; UNIV CALIFORNIA.		
DR	WPI: 98-077180/of; Richard V, Gorin F, Van PEL A;		

Page 169

```

FT      /'tag= 5
exon    2802..2906
        /*tag= h
exon    3031..3246
        /*tag= i

PD      W09120217-1A.
PP       14-OCT-1993.
PR       01-APR-1993. U03102.
PT       01-APR-1992. OS-961458
PI       01-APR-1992. OS-961458
PL       Johnson CD, Marchionni WA.
PW1:   93-336943/42.
DR       P-FSDM; R2742.
DT       Used to identify cloning of genes from lower organisms -
          particularly those that codes for evolutionary conserved
          proteins.
PT       aminoacid sequence.
PS       Disclosure; Fig 8; 18pp; English.
SC       The Primers/probes (Q426-Q4295) are used to isolate the ced-4
CC       gene from various species.
CC       Sequence 4093 BP; 1236 A; 752 C; 726 G; 1346 T;
          Score 14; DB 9; Length 4093;
          Local Similarity 10%; Pct. ID: 44%+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      1937 tgcatacacagca 1950
Cp      181 tacaacacacaga 168

RESULT 479 standard; DNA, 4100 BP.
Cc      QD 74052.
AC      Q47062.
Dt      28-JAN-1996 (first entry)
DE      The rat beta-actin gene.
DI      Rat.
OS      Beta-actin; primer; mRNA, specificity: pharmaceutical; aa.
CC      J07123984-A.
CC      16-MAY-1995.
PD      05-NOV-1993.
PP      05-NOV-1993.
PR      05-NOV-1993.
PT      05-NOV-1993.
PI      (HTEB.) HITACHI CHEM CO LTD.
PW1:   95-211627/28.
DR      Example 27; Page 28-39; 35pp; Japanese.
CC      Q74062 as the rat beta-actin gene. This gene is amplified by the
CC      primers Q74045 and Q74046. The primers are used specifically for the
CC      highest resolution detection of the beta-actin gene. They have the advantage of
CC      higher sensitivity than other methods.

```



CC Industry.  
SQ Sequence 4100 BP; 765 A; 1111 C; 1136 G; 1097 T;  
Query Match 1.38; Score 14; DB 16; Length 4100;  
Best Local Similarity 100.08; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1798 agctagctccct 1811  
|||||  
QY 1104 AGCTAGCTCCCT 1117

RESULT 480  
ID T58545 standard; cDNA; 4134 BP.  
AC T58545;  
DT 25-MAR-1997 (first entry) 3-kinase PI3K-gamma cDNA.  
KW Phosphatidylinositol 3-kinase; PI3K; gamma; immunosay;  
KW histamine secretion; nerve cell differentiation; glucose transport;  
KW modulation; regulation; Alzheimer's disease; lipolysis; ds.  
KW Key words sapiens. Location/Qualifiers  
FT cds  
FT 423..3572  
FT /\*tag= a  
FT /product= PI3K-gamma  
PD 04-APR-1996;  
PF 20-DEC-1994; 445562.  
PR 13-OCT-1994; DE-436696.  
PT PLAC 7 MAX PLANKER GES FORDERUNG WISSENSCHAFTEN.  
PW WPI: 96-172545/18.  
DR P-PDB: W1576.  
DT New phosphatidylinositol 3-kinase protein - useful as immunogen and  
DT of terms of kinase activity  
DT CDS 5' end: 100% GC content  
CC A 402 bp cDNA fragment was amplified from a human bone marrow  
CC library using PCR primers corresponding to amino acid sequences  
CC KMDPR and HDPG. The amplified fragment was used to probe a human  
CC cDNA library and several overlapping clones were isolated.  
CC The 1009 residues. The protein is a novel phosphatidylinositol 3-  
CC kinase (PI3K) that differs in its regulatory mechanism from the  
CC known PI3K-alpha and -beta enzymes. The new enzyme has been  
CC identified as a novel protein that can be used as an immunogen. The enzyme,  
CC antibodies against it, and the cDNA fragment can be used as an immunogen.  
CC modulating cell proliferation, receptor-mediated signal transmission,  
CC histamine secretion, nerve cell differentiation, glucose transport,  
CC and anti-lipolytic activity or for treating Alzheimer's disease.  
CC N.B. Although the claimed sequences are referred to by SEQ ID.

CC Numbers, a sequence listing did not appear in the original printed  
CC patent application.  
SQ Sequence 4134 BP; 1127 A; 999 C; 999 G; 1010 T;  
Query Match 1.38; Score 14; DB 26; Length 4134;  
Best Local Similarity 100.08; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 3079 ttcagcagctg 3092  
|||||  
Cp 908 TTCAGCAGCTG 895

RESULT 481  
ID T58545 standard; cDNA; 4137 BP.  
AC T58545;  
DT 25-MAR-1997 (first entry) 3-kinase PI3K-gamma cDNA.  
KW Phosphatidylinositol 3-kinase; PI3K; gamma; immunosay;  
KW histamine secretion; nerve cell differentiation; glucose transport;  
KW modulation; regulation; Alzheimer's disease; lipolysis; ds.  
KW Key words sapiens. Location/Qualifiers  
FT cds  
FT 423..3575  
FT /\*tag= a  
FT /product= PI3K-gamma  
PD 04-APR-1996;  
PF 20-DEC-1994; 445562.  
PR 13-OCT-1994; DE-436696.  
PT PLAC 7 MAX PLANKER GES FORDERUNG WISSENSCHAFTEN.  
PW WPI: 96-172545/18.  
DR P-PDB: W1577.  
DT New phosphatidylinositol 3-kinase protein - useful as immunogen and  
DT of terms of kinase activity  
DT CDS 5' end: 100% GC content  
CC A 402 bp cDNA fragment was amplified from a human bone marrow  
CC library using PCR primers corresponding to amino acid sequences  
CC KMDPR and HDPG. The amplified fragment was used to probe a human  
CC cDNA library and several overlapping clones were isolated.  
CC The 1009 residues. The protein is a novel phosphatidylinositol 3-  
CC kinase (PI3K) that differs in its regulatory mechanism from the  
CC known PI3K-alpha and -beta enzymes. The new enzyme has been  
CC identified as a novel protein that can be used as an immunogen. The enzyme,  
CC antibodies against it, and the cDNA fragment can be used as an immunogen.  
CC modulating cell proliferation, receptor-mediated signal transmission,  
CC histamine secretion, nerve cell differentiation, glucose transport,  
CC and anti-lipolytic activity or for treating Alzheimer's disease.  
CC N.B. Although the claimed sequences are referred to by SEQ ID.

CC Glucose transport and anti-lipolytic activity or for treating  
CC Alzheimer's disease.  
CC N.B. Although the claimed sequences are referred to by SEQ ID.  
CC Numbers, a sequence listing did not appear in the original printed  
CC patent application.  
SQ Sequence 4137 BP; 1127 A; 999 C; 1001 G; 1010 T;  
Query Match 1.38; Score 14; DB 26; Length 4137;  
Best Local Similarity 100.08; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 3082 ttcagcagctg 3095  
|||||  
Cp 908 TTCAGCAGCTG 895

RESULT 482  
ID Q2159 standard; DNA; 4155 BP.  
AC Q2159;  
DT 20-APR-1993 (first entry)  
DE BT toxin 17a.  
KW nematode worms; nematocidal toxin; agriculture; plants;  
KW Scopes; pests; Cry proteins.  
KW Bacillus thuringiensis.  
FT Key  
FT Location/Qualifiers  
FT cds  
FT 1..4155  
FT /\*tag= a  
PD 09-DEC-1992;  
PF 01-MAY-1992; 303969.  
PR 03-MAY-1991; US-693018.  
PR 31-JAN-1992; US-810050.  
PR 13-SEP-1991; US-753218.  
PT (MTCO) MYCOGEN CORP.  
PW Bagley AL, Cannon RJC, Payne JM;  
P-PDB: R3889/48.  
DT New Bacillus thuringiensis - useful for  
DT of animal or plant parasites; deoxyribonucleic acid  
DT control of animal or plant parasites; deoxyribonucleic acid  
DT Claim 1(b) : Page 23; 57pp; English.  
KW sequence encodes the Bacillus thuringiensis delta-endotoxin  
CC having sequence: transformed hosts and transgenic plants  
CC lambdaGEM11 using BT strain P817 total cellular DNA partially  
CC digested with SalI and size fractionated by electrophoresis  
CC (9-23kb used). Packaged phage were plated on E. coli RW51 and  
CC reselected, and positive plaques were purified and reselected  
CC separated by electrophoresis and reprobed. Two positive  
CC bands were present, clones contg. 4.5 and 2.7 kb EcoRI fragments.

CC These were cloned into pBlac, as E. coli/MT shuttle vector made up  
CC of replication origins from pUC18 and pUC19. This was used to  
CC transform E. coli MG21, and this grown on IPTG and XGAL media.  
CC White colonies were selected, purified, and either pUC1827 contg.  
CC the 2 kb P817a toxin fragment (this sequence) or pUC1828 contg.  
CC the 4 kb P817a toxin fragment (this sequence) and sequenced.  
SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;  
Query Match 1.38; Score 14; DB 6; Length 4155;  
Best Local Similarity 100.08; Pred. No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 402 aacttctactg 415  
|||||  
Cp 355 MCTTACTCTCCG 368

RESULT 483  
ID Q3083 standard; DNA; 4155 BP.  
AC Q3083;  
DT 01-APR-1993 (first entry)  
DE Toxin 17a.  
KW Endotoxin; acarides; pest; Two Spotted Spider; mite;  
KW Phytogagrus, sp.  
KW Bacillus thuringiensis P817a.  
PD 12-NOV-1992;  
PF 30-APR-1992; 003546.  
PR 03-MAY-1991; US-693018.  
PR 13-SEP-1991; US-753218.  
PT (MTCO) MYCOGEN CORP.  
PW Bagley AL, Cannon RJC, Payne JM;  
P-PDB: R3889/48.  
DT New Bacillus thuringiensis isolates and toxins - used for  
DT controlling acarid pests of livestock, fowl, stored prods. and  
DT plants.  
KW sequence: Page 25-27; 62pp; English.  
CC Gene sequences encoding a toxin which is active against acarides and  
CC is obtainable from B. thuringiensis isolates P817a, P817b, 3322,  
CC P852a1, P869D1, P868a1 and P850C are given in Q30803-07 and Q30820-21  
CC respectively. The genes encode a delta-endotoxin active against acarid  
CC pests including the Two Spotted Spider mite and the European spruce  
CC used against non-phytophagous mites such as acarid pests of livestock,  
CC fowl and stored prods. The genes can be cloned and used to  
CC transform other hosts, which can be used to control mites, or in  
CC the production of transgenic animals, be resistant to acarid pests.  
SQ Sequence 4155 BP; 1386 A; 668 C; 819 G; 1280 T;  
Query Match 1.38; Score 14; DB 5; Length 4155;



Best Local Similarity 100.0%; Pred. No. 3.47e+02; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttactgcgg 415

QY 355 AACCTTACTGCGG 368

RESULT 484

ID Q51687 standard; DNA; 4155 BP.  
AC Q51687-1995 (first entry)  
DE Bacillus thuringiensis crystal protein 17A gene.  
KW Crystal protein: delta endotoxin; acariicide; pesticide;  
KW biological control agent; transgenic plant; crop improvement; da.  
KW Bacillus thuringiensis isolate PS17a.  
PI Mat\_peptide 1: 4155  
FT /tag= a  
PP US262158-A.  
PP 10-APR-1991: 693210.  
PP 30-APR-1991: US-693210.  
PP 13-SEP-1991: US-759248.  
PP 30-SEP-1991: US-761841.  
PP 10-SEP-1991: US-761841.  
PA (MCO ) MYCOGEN CORP.  
PI Bagley AL, Cannon RJC, Payne JM;  
DR WPI: 93-37787/47.  
DR P-PSDB: R47801.  
PT Bacillus thuringiensis delta toxin cloned into microbe hosts.  
PT Dictionnaire; Column 21-26: 42pp; English.  
CC DNA encoding the insecticidal toxin can be cloned into baculo  
CC virus vectors. The host cells are preferably E. coli  
CC strains which become resistant to the acariicide pests.  
CC Specifically, the two-spotted spider mite (Tetranychus urticae) is  
CC controlled. 4155 BP; 1386 A; 677 C; 811 G; 1281 T;  
SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match 1.3%; Score 14; DB 15; Length 4155;

Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttactgcgg 415

QY 355 AACCTTACTGCGG 368

RESULT 485

nmatoxide worms; nematocidal; nematocidal toxin; agriculture; plants;

KW crops; pests; Cry proteins.

OS Bacillus Thuringiensis.

PI Mat\_peptide 1: 4155

FT /tag= a

PP US262158-A.

PP 10-APR-1991: 693210.

PP 30-APR-1991: US-693210.

PP 13-SEP-1991: US-759248.

PP 30-SEP-1991: US-761841.

PA (MCO ) MYCOGEN CORP.

PI Bagley AL, Cannon RJC, Payne JM, Schnepp HE, Schwab GE;

DR WPI: 93-39886/48.

DR P-PSDB: R28803.

PT New genes and toxins against nematodes - obtd. from Bacillus

PT thuringiensis strains with nematocidal activity

PT Dictionnaire; Column 21-26: 42pp; English.

CC This sequence encodes the Bacillus thuringiensis delta-endotoxin

CC having nematocidal activity. A library was constructed in

CC lambdaDE3 using Bt strain PS17 total cellular DNA partially

CC digested with BstI and ligated into lambdaDE3. The library

CC (9.2kb) was packaged into phage, test plated on E. coli NM522

CC and screened with probe Q30949. Positive plaques were purified and

CC rechecked, and positive plaques used to infect E. coli NM522 for

CC phage DNA preparation. Phage DNA was digested with EcoRI and

CC bands were present clones congs. 4.5 and 2.7 kb. EcoRI fragments

CC bands were cloned into pBCLac, an E. coli/Bst shuttle vector made up

CC of replication origins from pBCLac and pUC19. This was used to

CC transform E. coli NM522, and this grew on IPTG and XGAL media.

CC The 2.7kb PS17a toxin fragment (this sequence) or pBCLac28 conts.

CC the 4.5kb PS17b toxin fragment isolated and sequenced.

CC Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match 1.3%; Score 14; DB 15; Length 4155;

Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttactgcgg 415

QY 355 AACCTTACTGCGG 368

RESULT 487

AC Q50453 standard; DNA; 4155 BP.

AC Q50453-1996 (first entry)

DE PS17a acariide-active toxin DNA from strain PS17.

760069 standard; DNA; 4155 BP.

AC 760069-1997 (first entry)

DE 17A toxin coding sequence.

KW Toxin; ant; Bacillus thuringiensis; hymenopter pest; pharaoh ant;

KW biological control; Monomorium pharaonis; delta-endotoxin; lepidoptera;

OS Bacillus thuringiensis isolate PS17a.

PI Mat\_peptide 1: 4155

FT /tag= a

PP US559607-A.

PP 21-JAN-1997: 703977.

PP 22-MAY-1991: 703977.

PP 25-NOV-1991: US-703645.

PP 22-MAY-1992: US-887980.

PP 24-NOV-1993: US-158232.

PA (MCO ) MYCOGEN CORP.

PI Bagley AL, Cannon RJC, Payne JM;

DR WPI: 97-107615/10.

DR P-PSDB: W13884.

PT Bacillus thuringiensis toxin - active against hymenopter pests

PT Dictionnaire; Column 21-26: 42pp; English.

CC This sequence represents the coding sequence for the 17A toxin isolated

CC from the Bacillus thuringiensis (B.t.) isolate PS17a. B.t. is a

CC gram-positive, spore forming, soil bacterium, characterized by parasporal

CC crystalline protein inclusions. These proteins can be highly toxic to

CC lepidopteran insects. The proteins of the invention are active against

CC previously isolated B.t. delta-endotoxins were mainly active against

CC lepidopteran insects. However, the proteins of the invention are active

CC against hymenopter insects. The protein encoded by this sequence is an

CC insecticide active against the pest for which the sequence shown in

CC W13884 and W13871 represent the coding sequence for the toxin. The

CC invention are active against hymenopter pests, they can be used for the

CC biological control of ants, particularly pharaoh ants (Monomorium

CC pharaonis). 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match 1.3%; Score 14; DB 29; Length 4155;

Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttactgcgg 415

QY 355 AACCTTACTGCGG 368

Insecticide; toxin; acariide-active toxin; delta-endotoxin; PS17a; PS17;

OS two spotted spider mite; tetranychus urticae; Bacillus thuringiensis; da.

PI Mat\_peptide 1: 4155

FT /tag= a

PP US542410-A.

PP 10-APR-1991: 693210.

PP 30-APR-1991: US-693210.

PP 13-SEP-1991: US-759248.

PP 30-SEP-1991: US-761841.

PA (MCO ) MYCOGEN CORP.

PI Bagley AL, Cannon RJC, Payne JM;

DR WPI: 95-22377/29.

DR P-PSDB: W13884.

PT DNA encoding delta endotoxin of Bacillus thuringiensis - used for

PT the control of acarid pests and for the production of acarid pest

PT resistant plants

PT Dictionnaire; Column 21-26: 42pp; English.

CC This sequence encodes the Bacillus thuringiensis delta-endotoxin

CC having nematocidal activity. A library was constructed in

CC lambdaDE3 using Bt strain PS17 total cellular DNA partially

CC digested with BstI and ligated into lambdaDE3. The library

CC (9.2kb) was packaged into phage, test plated on E. coli NM522

CC and screened with probe Q30949. Positive plaques were purified and

CC rechecked, and positive plaques used to infect E. coli NM522 for

CC phage DNA preparation. Phage DNA was digested with EcoRI and

CC bands were present clones congs. 4.5 and 2.7 kb. EcoRI fragments

CC bands were cloned into pBCLac, an E. coli/Bst shuttle vector made up

CC of replication origins from pBCLac and pUC19. This was used to

CC transform E. coli NM522, and this grew on IPTG and XGAL media.

CC The 2.7kb PS17a toxin fragment (this sequence) or pBCLac28 conts.

CC the 4.5kb PS17b toxin fragment isolated and sequenced.

CC Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match 1.3%; Score 14; DB 16; Length 4155;

Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttactgcgg 415

QY 355 AACCTTACTGCGG 368

RESULT 488

ID Q72179 standard; DNA; 4155 BP.

AC Q72179-1996 (first entry)

DE PS17a acariide-active toxin DNA from strain PS17.

nmatoxide worms; nematocidal; nematocidal toxin; agriculture; plants;

KW crops; pests; Cry proteins.

OS Bacillus Thuringiensis.

PI Mat\_peptide 1: 4155

FT /tag= a

PP US262158-A.

PP 10-APR-1991: 693210.

PP 30-APR-1991: US-693210.

PP 13-SEP-1991: US-759248.

PP 30-SEP-1991: US-761841.

PA (MCO ) MYCOGEN CORP.

PI Bagley AL, Cannon RJC, Payne JM, Schnepp HE, Schwab GE;

DR WPI: 93-39886/48.

DR P-PSDB: R28803.

PT New genes and toxins against nematodes - obtd. from Bacillus

PT thuringiensis strains with nematocidal activity

PT Dictionnaire; Column 21-26: 42pp; English.

CC This sequence encodes the Bacillus thuringiensis delta-endotoxin

CC having nematocidal activity. A library was constructed in

CC lambdaDE3 using Bt strain PS17 total cellular DNA partially

CC digested with BstI and ligated into lambdaDE3. The library

CC (9.2kb) was packaged into phage, test plated on E. coli NM522

CC and screened with probe Q30949. Positive plaques were purified and

CC rechecked, and positive plaques used to infect E. coli NM522 for

CC phage DNA preparation. Phage DNA was digested with EcoRI and

CC bands were present clones congs. 4.5 and 2.7 kb. EcoRI fragments

CC bands were cloned into pBCLac, an E. coli/Bst shuttle vector made up

CC of replication origins from pBCLac and pUC19. This was used to

CC transform E. coli NM522, and this grew on IPTG and XGAL media.

CC The 2.7kb PS17a toxin fragment (this sequence) or pBCLac28 conts.

CC the 4.5kb PS17b toxin fragment isolated and sequenced.

CC Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

SQ Sequence 4155 BP; 1386 A; 677 C; 811 G; 1281 T;

Query Match 1.3%; Score 14; DB 15; Length 4155;

Best Local Similarity 100.0%; Pred. No. 3.47e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 402 aacttactgcgg 415

QY 355 AACCTTACTGCGG 368

RESULT 487

AC Q50453 standard; DNA; 4155 BP.

AC Q50453-1996 (first entry)

DE PS17a acariide-active toxin DNA from strain PS17.











PN EP-52922-A.  
 PP 27-AUG-1992; 114637  
 PR 29-AUG-1991; US-751460.  
 PA (AMCY ) AMERICAN CYANAMID CO.  
 PI Ishikawa Y., 168/09.  
 DR P-PSDB: R32882.  
 CC Isolated nucleic acid mol. encoding cardiac adenylyl cyclase type  
 PT v - useful for determining and modifying cardiac function  
 CC Claim 1; Page 15-27; 38pp; English.  
 CC A cDNA library was prepared in lambda gt10 phage. A 970 bp cDNA fragment  
 CC from type I adenylyl cyclase cDNA was used as probe. The  
 CC clones isolated were used to obtain cDNA encoding CACV. This probe  
 CC may also be used to screen a human cardiac cDNA library to obtain  
 CC cDNA encoding CACV. The cDNA encoding CACV may be used in  
 CC determining cardiac function. A decrease in CACV content of the  
 CC heart contributes to impaired cAMP prodn. and in heart failure. The  
 CC cDNA encoding CACV may be used to screen for cDNAs which stimulate or inhibit  
 CC the activity of CACV.  
 CC Sequence 4356 BP; 1367 G; 1377 G; 775 T;  
 SQ Query Match 1.3%; Score 14; DB 6; Length 4356;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1809 caagagacagcta 1822  
 Cp 86 CAGAGACAGCTA 73

RESULT 497  
 ID Q95440 standard; cDNA; 4356 BP.  
 DT 31-JAN-1996 (first entry)  
 DE Cardiac adenylyl cyclase gene.  
 KW Cardiac adenylyl cyclase; effector enzyme; ss.  
 PS Homo sapiens. Location/Qualifiers  
 FT cds 148..3702  
 FT /tag= a

PN TM-243453-A.  
 PP 02-JUN-1992; 105242  
 PR 12-JUN-1991; US-899068.  
 PA (AMCY ) AMERICAN CYANAMID CO.  
 DR PPI: 95-214006/28.  
 CC Cardiac adenylyl cyclase and corresp. DNA - having specified

PT sequences  
 CC Claim 1; Fig 2; 45pp; Chinese.  
 CC Q95440 encodes R78519, the novel effector enzyme cardiac adenylyl  
 CC cyclase.  
 CC Sequence 4356 BP; 836 A; 1373 C; 1373 G; 775 T;  
 SQ Query Match 1.3%; Score 14; DB 16; Length 4356;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1809 caagagacagcta 1822  
 Cp 86 CAGAGACAGCTA 73

RESULT 498  
 ID Q95440 standard; cDNA to mRNA; 4359 BP.  
 DT 20-NOV-1995 (first entry)  
 DE Brush-1 cDNA.  
 KW Brush-1; tumor suppressor gene; breast cancer; mamma carcinoma;  
 OS Homo sapiens. gene therapy; ss.  
 PH M09515334-A.  
 PP 08-JUN-1995.  
 PP 30-NOV-1994; U13823.  
 PR 27-SEP-1994; US-514598.  
 PA (CALP ) CALIFORNIA PACIFIC MEDICAL CENT RES INST.  
 PI Chang J., Schott DR. Smith RS;  
 DR WP1: 95-215228/28. from the Brush-1 tumour suppressor gene - useful  
 CC as probes for detecting tumours and premalignant cells, esp. of the  
 CC breast, and therapeutically  
 CC Claim 2; Page 25-29; 39pp; English.  
 CC cDNA obtd. from breast tumor cell RNA was subjected to PCR using  
 CC primers flanking the brush-1 gene. The resulting cDNA was sequenced  
 CC and compared with the brush-1 cDNA sequence. Isolated clones were used to assemble a Brush-1 cDNA  
 CC sequence that can be used to detect this marker in breast tissue.  
 CC Brush-1 nucleic acid may also be used for gene therapy in a  
 CC subject deficient in a functional Brush-1 gene.  
 CC Sequence 4359 BP; 1179 A; 352 G; 1373 T;  
 SQ Query Match 1.3%; Score 14; DB 15; Length 4359;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 4062 aaagtgtttgag 4075  
 Cp 1042 AAAGTGTTCGAC 1029

RESULT 499  
 ID Q61730 standard; cDNA; 4364 BP.  
 DT 17-OCT-1994 (first entry)  
 DE Protein tyrosine kinase JIL.  
 KW Protein tyrosine kinase; PTK; receptor-type protein-kinase; RTK;  
 OS Homo sapiens. Location/Qualifiers  
 FT Key 16..3384  
 FT cds 16..3384  
 FT /tag= a

PN W00410197-A.  
 PP 13-MAY-1994.  
 PR 30-OCT-1993; AU00581.  
 PA (LONM-) LUDWIG INST CANCER RES.  
 DR WP109-167386/20.  
 CC New protein tyrosine kinase, corresp. DNA and ligands - useful as  
 CC anticancer agents in treatment of carcinoma  
 CC Disclosure: Fig 1; 60pp; English.  
 CC The cDNA encoding the protein tyrosine kinase JIL was cloned from a  
 CC that was amplified by PCR. Chinese hamster ovary (CHO) cells  
 CC clone was used to screen a mouse lung cDNA library. Overlapping  
 CC clones were sequenced to generate the sequence given in Q61730.  
 CC JIL clone encodes a new protein tyrosine kinase that includes an  
 CC repeats an immunoglobulin-like domain and an epidermal growth  
 CC factor-like domain.  
 CC Sequence 4364 BP; 1257 A; 965 G; 1056 G; 1086 T;  
 SQ Query Match 1.3%; Score 14; DB 10; Length 4364;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 2694 agcactctgacc 2707  
 Qy 108 AGCCTATTCTGACC 121

RESULT 500  
 ID Q25444 standard; cDNA; 4405 BP.  
 AC Q25444;  
 DT 04-DEC-1992 (first entry)  
 DE Paired base-pairing sequence  
 KW Paired base-pairing sequence  
 OS Homo sapiens. Location/Qualifiers  
 FT Key 408..2795  
 FT cds 408..2795  
 FT /tag= a

PT signal\_peptide 408..485  
 PT /label= PACE  
 PT /note= includes two termination codons  
 PT /tag= b  
 PT poly\_a\_signal 418..4320  
 PT /tag= a  
 PT /note= possible polyA signal  
 PN W09205698-A.  
 PP 11-JUN-1992. U08725.  
 PR 26-NOV-1990; US-621092.  
 PR 29-NOV-1990; US-620859.  
 PR 29-NOV-1990; US-621443.  
 PR 30-NOV-1990; US-621457.  
 PA (GENY ) GENEMICS INST INC.  
 PI Barr PJ, Brake AJ, Kaufman RJ, Tekamp-Olsen P, Wasley L;  
 DR WP109-167386/20.  
 CC The coding sequence contains three consensus sites for N-linked  
 CC glycosylation and twenty-two cysteine residues. The active site is  
 CC in the ORF and includes a triad of amino acids: Asp153, His194 and  
 CC Ser368. Cysteine-rich region (CSR) is also present and is located in the  
 CC vicinity of amino acid Cys587 to amino acid Cys675. A putative  
 CC hydrophobic transmembrane domain (TM) is located downstream from the  
 CC cysteine-rich region, at ca. amino acid Val716 to amino acid Leu738.  
 CC Sequence 4405 BP; 822 A; 1442 C; 1330 G; 811 T;  
 SQ Query Match 1.3%; Score 14; DB 4; Length 4405;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 3437 ctgaacactctgact 3450  
 Qy 904 CTGAACACTCTGACT 917



Tue Nov 17 08:55:26 1998

RESULT	501	
ID	T05556 standard: DNA, 4405 BP.	
AC	T05556;	
CD	14-APR-1996 (first access)	
DE	Human PACE coding sequence.	
KE	Human PACE coding sequence.	
NC	subtilisin-like protease, precursor polypeptide processing;	
OS	blood coagulation protein precursor: ds.	
KS	Homo sapiens.	
WU		
SV	key	
FT	5'utr	Location/Qualifiers
FT	start	start
FT	signal_peptide	/*cgc a
FT	/*cgc b	408..485
FT	/*cgc c	485..792
FT	cds	485..792
FT	mat_peptide	/*cgc d
FT	485..798	485..798
FT	/*cgc e	/*cgc d
FT	/*cgc f	/*product: PACE
FT	3'utr	/*cgc g
FT	polya_signal	4345..4350
FT	4345..4350	/*cgc h
FT	/*cgc i	/*cgc f
FT	435460950..A.	
PN	26-NOV-1992: 885972.	
PP	20-MAY-1992: 885972.	
PP	26-NOV-1992: US-621092.	
PP	28-NOV-1990: US-621443.	
PP	28-NOV-1990: US-621443.	
PP	30-NOV-1990: US-621092.	
PP	30-NOV-1990: US-621092.	
PP	20-MAY-1992: US-885972.	
PP	(CHIR ) CHIRON CORP.	
PA	(SEMI ) GENETICS INST. INC.	
PI	Weng PA, Blake AJ, Kaufman RJ, Tekamp-Olson P, Wasley L;	
PI	WZ1: 95-373215/48.	
DR	P-PDSB: R75140.	
DR	Host cells expressing PACE for processing/conversion of precursor	
FT	biological activity and blood coagulation proteins.	
FT	biological activity and blood coagulation proteins.	
PP	Claim 7, Fig 2A-3F: 50pp: English.	
CC	The DNA encodes PACE, a paired basic amino acid converting enzyme,	
CC	which has structural homology with epsilon subtilisin, life serine	
CC	expressed in a host cell (yeast, mammalian or insect, preferably	
CC	CHO cells) along with a heterologous polynucleotide encoding a	
CC	precursor polypeptide preferably a precursor of a blood coagulation	
CC	activity, a heterologously expressed proteinase associated with	
CC	of basic amino acids (LeuArg, ArgArg, ArgArg) by the co-expressed	

Tue Nov 17 08:55:26 1998 US-08-887-977-9.mhq

```

DE KW Protein tyrosine-kinase SAL-S1 gene.
KW Protein tyrosine-kinase; PTK; SAL-S1; agonist; cell growth;
OS Homo sapiens.
NCBI differentiations: ss.
FT CD Location/Qualifiers
FT CDS 30..3977
FT /locus_tag="Sal-S1"
FT /db_xref="GeneID:147"
FT /db_xref="RefSeq:NM_001001"
FT /tag="a"
FT FT signal_peptide 30..102
FT /cds_tag="b"
FT /cds_tag="c"
FT FT mat_peptide 103..3977
FT /tag="a"
FT /tag="c"
FT WP0327061.1.
PD 13-OCT-1995.
PD 04-APR-1995.
PD 04-APR-1995.
PD (GENE) GENIECODE INC.
PD Bennett BD, Goeddel D, Lee JN, Matthews W, Tsai SP;
PD Wood WJ. 196160/47.
PD P-CDM8 R8937.
PD P-CDM8 R8937.
FT Agonist antibodies which activate specific protein tyrosine
FT kinase(s) - also activate chimeric proteins of kinase extracellular
FT domain and Ig constant domain, useful for studying, and therapeutic
FT applications of protein tyrosine kinase.
FT Description of this protein:
FT PTK domain of 1315-1725 amino acids.
FT DNA probes cDNA libraries to identify novel PTK genes.
FT The SAL-S1 gene (see also 703101) was isolated from several megakaryocytic cell
FT lines. The gene was used to produce recombinant SAL-S1 or its
FT fragments. These genes were used to produce antibodies, oligonucleotides
FT or antisense nucleotides that modulate PTK activity.
FT Sequence 4425 BP; 939 A; 1348 C; 1361 G; 777 T;
FT Query Match 138; Score 14; DB 16; Length 4425;
FT Best Local Similarity 100.0%; P-CDM8 3,474/402;
FT Matches 0; Mismatches 0; Indels 0; Gaps 0;
DB 4051 cc05gagccgcaaga 4064
Cc 160 CCGAGGCGCGTACG 147

```

RESULT	504
ID	T89487 standard; cDNA; 4440 BP.
AC	T89487;
DT	19-FEB-1998 (first entry)
DE	Human A20 protein cDNA.
KW	A20; anti-apoptotic protein; human; nuclear factor-kappa B;
OS	Human carcinoma transplant;
	Inflammation; gene therapy; endothelial cell; dn.
	Homo sapiens.

[illegible]

Tue Nov 17 08:55:26 1998  
115-08-887-977-9

	Key	Location/Qualifiers	
	CDS	67-2439	
	WtSeq	/etseq = a	
DB	M09730083.AL.		
PF	21-AUG-1997		
PP	11-AUG-1997	E00676	
PR	19-APR-1996	US-634995	
PX	14-FEB-1996	US-601515	
RFA	(NEWS-) NEW ENGLAND DEACONESS HOSPITAL.		
RFN	(MOV) NOVARTIS AG.		
RT	NPI 97-424975/35		
SD	NPI 97-424975/35		
SR	P-PUBS; W31528		
ST	Recombinant endothelial cell containing DNA encoding anti-apoptotic protein is less susceptible to inflammatory response and is resistant to apoptosis after transplantation		
TT	Example 8; Page 10-44; 75pp; English		
CC	This DNA clone encodes human A30 (see W31528), a protein capable of blocking or suppressing NF-kappa B (NF- $\kappa$ B) activation. It was obtained from HUVEC cells by a claimed method of genetically modifying cells to express a desired gene product.		
CC	The protein encoded by this DNA clone is referred to as A30, which is also referred to as an inflammatory or other immunological stimulus comprising inserting into the cell, DNA encoding an anti-apoptotic protein able to inhibit NF- $\kappa$ B, and expressing the cell such that the regulation of the cell is limited in the presence of the stimulus.		
CC	The cell is transfected with a construct that includes at least one of the following: A30 (see W31528); BCL-XL (see W31530); and A1 (see W31531) and their deletion mutants. Also claimed are: (1) a mammalian endothelial cell modified by the above method; and (2) a recombinant transgenic or somatic recombinant animal comprising DNA encoding A30 or a deletion mutant thereof.		
CC	The above methods can be used to generate donor endothelial cells or graftable tissues or organs for transplantation into recipient species.		
CC	Sequence 4440 BP: 1206 A; 1070 C; 1055 G; 1109 T; 4440 BP: 1206 A; 1070 C; 1055 G; 1109 T;		
CC	Query Match		
CC	Best Local Similarity 100.0%; Pred. No. 8.37e+01; Length 4440;		
CC	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
DB	2483 gensecondat 2497		
DB	492 GCGTCACTCATCAT		

RESULT	505
ID	T64783
standard:	cDNA; 4503 BP.
	T64783;
	03-SEP-1997 (first entry)
	Human oxygen regulated protein ORP 150 cDNA.
	Oxygen regulated protein; ORP 150; stress protein; lachnaemia;
	hypoglycaemia; hypoxia; gene therapy; diabetes; A-







US-08-887-977-9-109

```

FT      /tag= c
FT intron 1725..2091
FT exon   /tag= d
FT       2092..2271
FT intron /tag= e
FT       2272..2375
FT exon   /tag= f
FT       2376..2673
FT intron /tag= g
FT       2674..2856
FT exon   /tag= h
FT       2857..2907
FT intron /tag= i
FT       2908..2994
FT exon   /tag= j
FT       2995..3251
FT intron /tag= k
FT       3252..3340
FT exon   /tag= l
FT       3341..3575
FT intron /tag= m
FT       3576..3670
FT exon   /tag= n
FT       3671..3792
FT       /tag= o
J02242687.A.
14-SEP-1989.    061702
14-MAR-1989; JP-061702.
(MOCH) MOCHIDA PHARM KK.
WT: 50..364257/49.
Plasmid and expression plasmid - contains promoter region of
human polypeptide chain elongation factor
Disclosure: Fig 3; Jppp; Japanese.
CC The nucleotide sequence may be used as an expression
CC plasmid in transgenic animals without limitation of territory
over plasmids using the SV40 promoter, stable for a month in cells,
and able to express a wide range of products.
Sequence 4691 BP; 1200 A; 990 C; 1233 G; 1268 T;
Query Match 1.3%; Score 14; DB 1; Length 4691;
Best Local Similarity 100.0%; Pred.No. 3.47e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps;
DB 2991 gtttcttatccc 2304
QY 637 GGTTCCTTTATCCC 650
RESULT 510

```

Page 505

[illegible]

Key	Location/Qualifiers
cds	/tag= a
Fr	/product= Capsular polysaccharide CpsD.
Fr	/tag= b
Fr	/product= Capsular polysaccharide CpsD.
Fr	/tag= b3
cds	/tag= b
Fr	/product= Capsular polysaccharide CpsD.
Fr	/tag= c
Fr	/product= Capsular polysaccharide CpsD.
Fr	/tag= c3
cds	/tag= c
Fr	/product= Capsular polysaccharide CpsD.
Fr	/tag= d3



PT	/product= Capsular polysaccharide CpsM.
PN	W0853548-A1.
PP	23-NOV-1995. U05138.
PD	16-MAY-1994. US-243546.
PR	(UABF) UAB RES FOUND.
PA	Dilled JF, 0317; Other J;
PI	Dilled JF, 0317; Other J;
PJ	N-PSDB: R3A0018, R3A0019, R3A0040, R3A0041.
PK	New Streptococcus pneumoniae capsular polysaccharide genes - used
PL	for detection, serotyping and for diagnosis and prevention of S.
PM	Sequences encoding the 5' flanking region of the capsular
PN	polysaccharide gene (cps), of Streptococcus pneumoniae and which are
PO	hybridized under conditions to S pneumoniae cps gene flanking region
PP	may be used in methods to detect and serotype S.pneumoniae. They
PQ	may also be used for the diagnosis and prevention of S. pneumoniae
PR	Sequence". 4951 BP; 1565 A; 749 C; 1028 G; 1609 T;
PS	Query Match Score 14: 100.0%; Length 4951;
PT	Query Local Similarity 100.0%; Query No. 0;
PU	Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps
PV	3032 Translating 3045
PW	17 TAAAGATTGTTTGG 30
PX	TAAAGATTGTTTGG
PY	TAAAGATTGTTTGG
PZ	TAAAGATTGTTTGG

RESULT	511
ID	G7345 standard; CDNA; 4975 BP.
AC	G73445;
DT	09-MAY-1995 (first entry)
DZ	IGF-I receptor;
DE	IGF-I receptor; insulin-like growth factor-I receptor;
OS	cell proliferation; Hela; mutagenesis; tyrosine-kinase;
KW	phosphorylation; probe; ds.
KS	Homo sapiens.
FT	Location/Qualifiers
FT	32..4135
FT	/tag= a
FT	signal_peptide
FT	32..121 b
FT	122..229 b
FT	mat_peptide
FT	/tag= c
FT	/product= IGF-I receptor alpha subunit
FT	2452..4132
FT	mat_peptide
FT	/product= IGF-I receptor beta subunit
FT	2452..4132

Tue Nov 17 08:55:26 1998  
US-08-887-977-9.FAQ

DB WPI: 96-259475/26.  
P-RSDB: R95244.  
PPT Induction of resistance to tumour cell growth - by insertion of  
diffusion chamber contg. tumour cell culture supplemented with, e.g.  
apoptosis inducing agent, into host.  
PS Diaclosole; Fig 4A-G; 47pp; English.  
CCC The nucleotide sequence (729608) coding for insulin growth factor  
RNA sequences (729606 and 729607) complementary to antisense 198 of  
the mature protein and codons -79 to -24 of the signal sequence  
respectively. The antisense sequences act as apoptosis-inducing  
agents when supplied to tumour cells growing in a diffusion chamber.  
CCC Induction of resistance to tumour growth occurs upon insertion  
of diffusion chamber into a mouse.  
DB Sequence 4989-587. 1369 C; 1322 G; 1082 T;  
580 Query Match 1.3%; Score 14; DB 20; Length 4989;  
Best Local Similarity 100.0%; Ref.No. 3.47e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps  
0  
DB 188 ccctggcaagacgc 201  
CP 111 ccctggcaagacgc 98  
|||||

[illegible]

Page 503

Thu Nov 17 08:55:26 1998

ATG-directed sense oligonucleotide(s) - useful for regulating transcription factor gene expression for e.g. wound healing and fibrosis

PP This is a patent application.

PS Disclosure: Page 46-52; 73pp; English.

PB A cDNA clone (Tg13129) codes for human insulin-like growth factor I receptor (IGF IR) (M91429).

CCC It plays a crucial role in

CCC vascular smooth muscle cell (VSMC) proliferative responses. Rat

CCC (see Tg13125 and Tg13128) to down-regulate IGF IR gene expression and

CCC retard VSMC growth, e.g. to treat restenosis and atherosclerosis.

CCC and also for ATP-directed sense oligonucleotide (see Tg13126),

CCC methods to stimulate VSMC growth e.g. for healing wounds and burns;

SC Sequence 4989 Ppt. 1116 A; 1168 C; 1121 G; 1082 T;

Query Match 1.3%; Score 14; DB 19; Length 4989;

Best Local Similarity 100.0%;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dbs 188 scgtsgagactgc 201

Cp 111 GCCTTGAGAACTGC 98

RESULT	516	
ID	T95673	standard; cDNA; 4989 BP.
AC	D	
CC	T95673;	
CD	14-APR-1998	(first entry)
DE	insulin-like growth factor I receptor cDNA	
EF	insulin-like growth factor I receptor cDNA	
EW	inhibitor; active survival domain; cell proliferation; cell death;	
KW	cancer; tumour; lymphoma; autoimmune disease; AIDS; inflammation;	
KW	therapy; human; ss.	
OR	human adipens.	
PF	Key	Location/Qualifiers
PF	CD5	46..4149
PF	sig_peptide	/start=a
PF	sig_peptide	/end=b
PF	mat_peptide	/start=
PF	mat_peptide	136..4146
PF		/end=c
PF	W9071010..11	
PF	09-OCT-1997	
PF	ED	
PF	01-APR-1997	U06087.
PF	01-APR-1996	U5-625819.
PF	(UNIQ)	W9071010..11
PF	(UNIQ)	W9071010..11
PF	Baserna RL, Jefferson THOMAS.	
P1	WPI; 57-503097/46.	
DR	PF=PB; W97692.	
DR	specifically C-terminally truncated or derived peptide(s), useful for	

Page 504



modulating apoptosis in e.g. cancer, AIDS, autoimmune diseases,  
PF Disclosure: Page 60-65; 110pp; English.  
PS This cDNA sequence codes for human insulin-like growth factor I  
CC receptor (IGF-IR, see W3762). Claimed compositions comprise one  
CC of: (a) a deletion mutant of IGF-IR which has part of the  
CC C-terminal 108-92 or 44 amino acids of IGF-IR; or (c) point  
CC mutants of IGF-IR, especially Y1251P, Y1250Y/Y1251P or H1293P/Y1294R.  
CC The peptides are derived from regions of the receptor which are  
CC essential for the proliferative and transforming activity of  
CC IGF-IR. The peptides are used in the treatment of cancer and to  
CC induce or suppress apoptosis in cells and/or tissues of an  
CC individual suffering from a disorder characterised by inappropriate  
CC cell proliferation or survival or cell death, such as inflammatory  
CC conditions, cancer including lymphomas and genotypic tumours,  
CC or cancer of the prostate, breast, colon, lung, stomach, pancreas,  
CC or chemotherapy or acute hypoxic injury, etc. Cytokine-dependent  
CC cells stably transfected with cDNA encoding IGF-IR are used in  
CC claimed methods for identifying agents that modulate apoptosis or  
CC which are capable of enhancing the anti-apoptotic response of  
CC various cells or tissues.  
SQ Sequence 4989 BP; 1316 A; 1371 C; 1320 G; 1082 T;  
Query Match 1.38; Score 14; DB 37; Length 4989;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0;  
Db 188 gccggagactgc 201  
Cp 111 GCCTGGAGACTGC 98

RESULT 517  
ID Q3702 standard; DNA; 4989 BP.  
DT 08-MAY-1995 (first entry)  
DE Human IGF-1 receptor.  
KW Insulin-like growth factor-1 receptor; IGF-IR; signal peptide;  
KW antisense; oligonucleotide; differentiation; cancer therapy;  
KW Inflammation; proliferation; cell; skin; keratinocyte growth factor;  
OS Homo sapiens. Location/Qualifiers  
FH Key 46..1354  
FI cds  
FI signal\_peptide 46..1354  
FI tag a  
FI tag b  
FI tag c  
FI tag d  
FI tag e  
FI tag f  
FI tag g  
FI tag h  
FI tag i  
FI tag j  
FI tag k  
FI tag l  
FI tag m  
FI tag n  
FI tag o  
FI tag p  
FI tag q  
FI tag r  
FI tag s  
FI tag t  
FI tag u  
FI tag v  
FI tag w  
FI tag x  
FI tag y  
FI tag z  
FI tag aa  
FI tag ab  
FI tag ac  
FI tag ad  
FI tag ae  
FI tag af  
FI tag ag  
FI tag ah  
FI tag ai  
FI tag aj  
FI tag ak  
FI tag al  
FI tag am  
FI tag an  
FI tag ao  
FI tag ap  
FI tag aq  
FI tag ar  
FI tag as  
FI tag at  
FI tag au  
FI tag av  
FI tag aw  
FI tag ax  
FI tag ay  
FI tag az  
FI tag ba  
FI tag bb  
FI tag bc  
FI tag bd  
FI tag be  
FI tag bf  
FI tag bg  
FI tag bh  
FI tag bi  
FI tag bj  
FI tag bk  
FI tag bl  
FI tag bm  
FI tag bn  
FI tag bo  
FI tag bp  
FI tag bq  
FI tag br  
FI tag bs  
FI tag bt  
FI tag bu  
FI tag bv  
FI tag bw  
FI tag bx  
FI tag by  
FI tag bz  
FI tag ca  
FI tag cb  
FI tag cc  
FI tag cd  
FI tag ce  
FI tag cf  
FI tag cg  
FI tag ch  
FI tag ci  
FI tag cj  
FI tag ck  
FI tag cl  
FI tag cm  
FI tag cn  
FI tag co  
FI tag cp  
FI tag cq  
FI tag cr  
FI tag cs  
FI tag ct  
FI tag cu  
FI tag cv  
FI tag cw  
FI tag cx  
FI tag cy  
FI tag cz  
FI tag da  
FI tag db  
FI tag dc  
FI tag dd  
FI tag de  
FI tag df  
FI tag dg  
FI tag dh  
FI tag di  
FI tag dj  
FI tag dk  
FI tag dl  
FI tag dm  
FI tag dn  
FI tag do  
FI tag dp  
FI tag dq  
FI tag dr  
FI tag ds  
FI tag dt  
FI tag du  
FI tag dv  
FI tag dw  
FI tag dx  
FI tag dy  
FI tag dz  
FI tag ea  
FI tag eb  
FI tag ec  
FI tag ed  
FI tag ee  
FI tag ef  
FI tag eg  
FI tag eh  
FI tag ei  
FI tag ej  
FI tag ek  
FI tag el  
FI tag em  
FI tag en  
FI tag eo  
FI tag ep  
FI tag eq  
FI tag er  
FI tag es  
FI tag et  
FI tag eu  
FI tag ev  
FI tag ew  
FI tag ex  
FI tag ey  
FI tag ez  
FI tag fa  
FI tag fb  
FI tag fc  
FI tag fd  
FI tag fe  
FI tag ff  
FI tag fg  
FI tag fh  
FI tag fi  
FI tag fj  
FI tag fk  
FI tag fl  
FI tag fm  
FI tag fn  
FI tag fo  
FI tag fp  
FI tag fq  
FI tag fr  
FI tag fs  
FI tag ft  
FI tag fu  
FI tag fv  
FI tag fw  
FI tag fx  
FI tag fy  
FI tag fz  
FI tag ga  
FI tag gb  
FI tag gc  
FI tag gd  
FI tag ge  
FI tag gf  
FI tag gg  
FI tag gh  
FI tag gi  
FI tag gj  
FI tag gk  
FI tag gl  
FI tag gm  
FI tag gn  
FI tag go  
FI tag gp  
FI tag gq  
FI tag gr  
FI tag gs  
FI tag gt  
FI tag gu  
FI tag gv  
FI tag gw  
FI tag gx  
FI tag gy  
FI tag gz  
FI tag ha  
FI tag hb  
FI tag hc  
FI tag hd  
FI tag he  
FI tag hf  
FI tag hg  
FI tag hh  
FI tag hi  
FI tag hj  
FI tag hk  
FI tag hl  
FI tag hm  
FI tag hn  
FI tag ho  
FI tag hp  
FI tag hq  
FI tag hr  
FI tag hs  
FI tag ht  
FI tag hu  
FI tag hv  
FI tag hw  
FI tag hx  
FI tag hy  
FI tag hz  
FI tag ia  
FI tag ib  
FI tag ic  
FI tag id  
FI tag ie  
FI tag if  
FI tag ig  
FI tag ih  
FI tag ii  
FI tag ij  
FI tag ik  
FI tag il  
FI tag im  
FI tag in  
FI tag io  
FI tag ip  
FI tag iq  
FI tag ir  
FI tag is  
FI tag it  
FI tag iu  
FI tag iv  
FI tag iw  
FI tag ix  
FI tag iy  
FI tag iz  
FI tag ja  
FI tag jb  
FI tag jc  
FI tag jd  
FI tag je  
FI tag jf  
FI tag jg  
FI tag jh  
FI tag ji  
FI tag jj  
FI tag jk  
FI tag jl  
FI tag jm  
FI tag jn  
FI tag jo  
FI tag jp  
FI tag jq  
FI tag jr  
FI tag js  
FI tag jt  
FI tag ju  
FI tag jv  
FI tag jw  
FI tag jx  
FI tag jy  
FI tag jz  
FI tag ka  
FI tag kb  
FI tag kc  
FI tag kd  
FI tag ke  
FI tag kf  
FI tag kg  
FI tag kh  
FI tag ki  
FI tag kj  
FI tag kk  
FI tag kl  
FI tag km  
FI tag kn  
FI tag ko  
FI tag kp  
FI tag kq  
FI tag kr  
FI tag ks  
FI tag kt  
FI tag ku  
FI tag kv  
FI tag kw  
FI tag kx  
FI tag ky  
FI tag kz  
FI tag la  
FI tag lb  
FI tag lc  
FI tag ld  
FI tag le  
FI tag lf  
FI tag lg  
FI tag lh  
FI tag li  
FI tag lj  
FI tag lk  
FI tag ll  
FI tag lm  
FI tag ln  
FI tag lo  
FI tag lp  
FI tag lq  
FI tag lr  
FI tag ls  
FI tag lt  
FI tag lu  
FI tag lv  
FI tag lw  
FI tag lx  
FI tag ly  
FI tag lz  
FI tag ma  
FI tag mb  
FI tag mc  
FI tag md  
FI tag me  
FI tag mf  
FI tag mg  
FI tag mh  
FI tag mi  
FI tag mj  
FI tag mk  
FI tag ml  
FI tag mn  
FI tag mo  
FI tag mp  
FI tag mq  
FI tag mr  
FI tag ms  
FI tag mt  
FI tag mu  
FI tag mv  
FI tag mw  
FI tag mx  
FI tag my  
FI tag mz  
FI tag na  
FI tag nb  
FI tag nc  
FI tag nd  
FI tag ne  
FI tag nf  
FI tag ng  
FI tag nh  
FI tag ni  
FI tag nj  
FI tag nk  
FI tag nl  
FI tag nm  
FI tag no  
FI tag np  
FI tag nq  
FI tag nr  
FI tag ns  
FI tag nt  
FI tag nu  
FI tag nv  
FI tag nw  
FI tag nx  
FI tag ny  
FI tag nz  
FI tag oa  
FI tag ob  
FI tag oc  
FI tag od  
FI tag oe  
FI tag of  
FI tag og  
FI tag oh  
FI tag oi  
FI tag oj  
FI tag ok  
FI tag ol  
FI tag om  
FI tag on  
FI tag oo  
FI tag op  
FI tag oq  
FI tag or  
FI tag os  
FI tag ot  
FI tag ou  
FI tag ov  
FI tag ow  
FI tag ox  
FI tag oy  
FI tag oz  
FI tag pa  
FI tag pb  
FI tag pc  
FI tag pd  
FI tag pe  
FI tag pf  
FI tag pg  
FI tag ph  
FI tag pi  
FI tag pj  
FI tag pk  
FI tag pl  
FI tag pm  
FI tag pn  
FI tag po  
FI tag pp  
FI tag pq  
FI tag pr  
FI tag ps  
FI tag pt  
FI tag pu  
FI tag pv  
FI tag pw  
FI tag px  
FI tag py  
FI tag pz  
FI tag qa  
FI tag qb  
FI tag qc  
FI tag qd  
FI tag qe  
FI tag qf  
FI tag qg  
FI tag qh  
FI tag qi  
FI tag qj  
FI tag qk  
FI tag ql  
FI tag qm  
FI tag qn  
FI tag qo  
FI tag qp  
FI tag qq  
FI tag qr  
FI tag qs  
FI tag qt  
FI tag qu  
FI tag qv  
FI tag qw  
FI tag qx  
FI tag qy  
FI tag qz  
FI tag ra  
FI tag rb  
FI tag rc  
FI tag rd  
FI tag re  
FI tag rf  
FI tag rg  
FI tag rh  
FI tag ri  
FI tag rj  
FI tag rk  
FI tag rl  
FI tag rm  
FI tag rn  
FI tag ro  
FI tag rp  
FI tag rq  
FI tag rr  
FI tag rs  
FI tag rt  
FI tag ru  
FI tag rv  
FI tag rw  
FI tag rx  
FI tag ry  
FI tag rz  
FI tag sa  
FI tag sb  
FI tag sc  
FI tag sd  
FI tag se  
FI tag sf  
FI tag sg  
FI tag sh  
FI tag si  
FI tag sj  
FI tag sk  
FI tag sl  
FI tag sm  
FI tag sn  
FI tag so  
FI tag sp  
FI tag sq  
FI tag sr  
FI tag ss  
FI tag st  
FI tag su  
FI tag sv  
FI tag sw  
FI tag sx  
FI tag sy  
FI tag sz  
FI tag ta  
FI tag tb  
FI tag tc  
FI tag td  
FI tag te  
FI tag tf  
FI tag tg  
FI tag th  
FI tag ti  
FI tag tj  
FI tag tk  
FI tag tl  
FI tag tm  
FI tag tn  
FI tag to  
FI tag tp  
FI tag tq  
FI tag tr  
FI tag ts  
FI tag tt  
FI tag tu  
FI tag tv  
FI tag tw  
FI tag tx  
FI tag ty  
FI tag tz  
FI tag ua  
FI tag ub  
FI tag uc  
FI tag ud  
FI tag ue  
FI tag uf  
FI tag ug  
FI tag uh  
FI tag ui  
FI tag uj  
FI tag uk  
FI tag ul  
FI tag um  
FI tag un  
FI tag uo  
FI tag up  
FI tag uq  
FI tag ur  
FI tag us  
FI tag ut  
FI tag uu  
FI tag uv  
FI tag vw  
FI tag vx  
FI tag vy  
FI tag vz  
FI tag wa  
FI tag wb  
FI tag wc  
FI tag wd  
FI tag we  
FI tag wf  
FI tag wg  
FI tag wh  
FI tag wi  
FI tag wj  
FI tag wk  
FI tag wl  
FI tag wm  
FI tag wn  
FI tag wo  
FI tag wp  
FI tag wq  
FI tag wr  
FI tag ws  
FI tag wt  
FI tag wu  
FI tag wv  
FI tag wx  
FI tag wy  
FI tag wz  
FI tag xa  
FI tag xb  
FI tag xc  
FI tag xd  
FI tag xe  
FI tag xf  
FI tag xg  
FI tag xh  
FI tag xi  
FI tag xj  
FI tag xk  
FI tag xl  
FI tag xm  
FI tag xn  
FI tag xo  
FI tag xp  
FI tag xq  
FI tag xr  
FI tag xs  
FI tag xt  
FI tag xu  
FI tag xv  
FI tag xw  
FI tag xx  
FI tag xy  
FI tag yz  
FI tag za  
FI tag zb  
FI tag zc  
FI tag zd  
FI tag ze  
FI tag zf  
FI tag zg  
FI tag zh  
FI tag zi  
FI tag zj  
FI tag zk  
FI tag zl  
FI tag zm  
FI tag zn  
FI tag zo  
FI tag zp  
FI tag zq  
FI tag zr  
FI tag zs  
FI tag zt  
FI tag zu  
FI tag zv  
FI tag zw  
FI tag zx  
FI tag zy  
FI tag zz

25-MAR-1994; U03114.  
PF 16-MAR-1993; US-021257.  
PS 16-MAR-1993; US-021257.  
PA (GDB:) UNIV JEFFERSON THOMAS.  
PI Baeciga R, Rubin R, Sell C.  
DR WPI: 94-332832/41.  
DE Inhibition of cell growth and differentiation of cells using anti-sense  
CC oligo-nucleotide(s) complementary to insulin-like growth factor  
CC 1 receptor; also causes differentiation and is used to treat  
CC cancer.  
CC This disclosure describes a method for the treatment of cancer by  
CC the administration of oligonucleotides complementary to the  
CC complete sequence is given in Q3702, and shows greater  
CC inhibition of the growth of human prostatic cancer, ovarian cancer  
CC and melanoma cell lines than the corresponding sense (Q3697)  
CC and antisense oligonucleotides (Q3698).  
SQ Sequence 4989 BP; 1316 A; 1371 C; 1320 G; 1082 T;  
Query Match 1.38; Score 14; DB 12; Length 4989;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0;  
Db 188 gccggagactgc 201  
Cp 111 GCCTGGAGACTGC 98

RESULT 518  
ID Q9245 standard; DNA; 4989 BP.  
DT 27-JAN-1997 (first entry)  
DE Insulin like growth factor-1 receptor.  
KW Insulin like growth factor; binding protein; IGF; IGFPR;  
KW Inflammation; proliferation; cell; skin; keratinocyte growth factor;  
OS Homo sapiens. Location/Qualifiers  
FH Key 32..4132  
FI cds  
FI signal\_peptide 32..121  
FI tag a IGF-1 receptor  
FI tag b  
FI tag c  
FI tag d  
FI tag e  
FI tag f  
FI tag g  
FI tag h  
FI tag i  
FI tag j  
FI tag k  
FI tag l  
FI tag m  
FI tag n  
FI tag o  
FI tag p  
FI tag q  
FI tag r  
FI tag s  
FI tag t  
FI tag u  
FI tag v  
FI tag w  
FI tag x  
FI tag y  
FI tag z  
FI tag aa  
FI tag ab  
FI tag ac  
FI tag ad  
FI tag ae  
FI tag af  
FI tag ag  
FI tag ah  
FI tag ai  
FI tag aj  
FI tag ak  
FI tag al  
FI tag am  
FI tag an  
FI tag ao  
FI tag ap  
FI tag aq  
FI tag ar  
FI tag as  
FI tag at  
FI tag au  
FI tag av  
FI tag aw  
FI tag ax  
FI tag ay  
FI tag az  
FI tag ba  
FI tag bb  
FI tag bc  
FI tag bd  
FI tag be  
FI tag bf  
FI tag bg  
FI tag bh  
FI tag bi  
FI tag bj  
FI tag bk  
FI tag bl  
FI tag bm  
FI tag bn  
FI tag bo  
FI tag bp  
FI tag bq  
FI tag br  
FI tag bs  
FI tag bt  
FI tag bu  
FI tag bv  
FI tag bw  
FI tag bx  
FI tag by  
FI tag bz  
FI tag ca  
FI tag cb  
FI tag cc  
FI tag cd  
FI tag ce  
FI tag cf  
FI tag cg  
FI tag ch  
FI tag ci  
FI tag cj  
FI tag ck  
FI tag cl  
FI tag cm  
FI tag cn  
FI tag co  
FI tag cp  
FI tag cq  
FI tag cr  
FI tag cs  
FI tag ct  
FI tag cu  
FI tag cv  
FI tag cw  
FI tag cx  
FI tag cy  
FI tag cz  
FI tag da  
FI tag db  
FI tag dc  
FI tag dd  
FI tag de  
FI tag df  
FI tag dg  
FI tag dh  
FI tag di  
FI tag dj  
FI tag dk  
FI tag dl  
FI tag dm  
FI tag dn  
FI tag do  
FI tag dp  
FI tag dq  
FI tag dr  
FI tag ds  
FI tag dt  
FI tag du  
FI tag dv  
FI tag dw  
FI tag dx  
FI tag dy  
FI tag dz  
FI tag ea  
FI tag eb  
FI tag ec  
FI tag ed  
FI tag ee  
FI tag ef  
FI tag eg  
FI tag eh  
FI tag ei  
FI tag ej  
FI tag ek  
FI tag el  
FI tag em  
FI tag en  
FI tag eo  
FI tag ep  
FI tag eq  
FI tag er  
FI tag es  
FI tag et  
FI tag eu  
FI tag ev  
FI tag ew  
FI tag ex  
FI tag ey  
FI tag ez  
FI tag fa  
FI tag fb  
FI tag fc  
FI tag fd  
FI tag fe  
FI tag ff  
FI tag fg  
FI tag fh  
FI tag fi  
FI tag fj  
FI tag fk  
FI tag fl  
FI tag fm  
FI tag fn  
FI tag fo  
FI tag fp  
FI tag fq  
FI tag fr  
FI tag fs  
FI tag ft  
FI tag fu  
FI tag fv  
FI tag fw  
FI tag fx  
FI tag fy  
FI tag fz  
FI tag ga  
FI tag gb  
FI tag gc  
FI tag gd  
FI tag ge  
FI tag gf  
FI tag gg  
FI tag gh  
FI tag gi  
FI tag gj  
FI tag gk  
FI tag gl  
FI tag gm  
FI tag gn  
FI tag go  
FI tag gp  
FI tag gq  
FI tag gr  
FI tag gs  
FI tag gt  
FI tag gu  
FI tag gv  
FI tag gw  
FI tag gx  
FI tag gy  
FI tag gz  
FI tag ha  
FI tag hb  
FI tag hc  
FI tag hd  
FI tag he  
FI tag hf  
FI tag hg  
FI tag hh  
FI tag hi  
FI tag hj  
FI tag hk  
FI tag hl  
FI tag hm  
FI tag hn  
FI tag ho  
FI tag hp  
FI tag hq  
FI tag hr  
FI tag hs  
FI tag ht  
FI tag hu  
FI tag hv  
FI tag hw  
FI tag hx  
FI tag hy  
FI tag hz  
FI tag ia  
FI tag ib  
FI tag ic  
FI tag id  
FI tag ie  
FI tag if  
FI tag ig  
FI tag ih  
FI tag ii  
FI tag ij  
FI tag ik  
FI tag il  
FI tag im  
FI tag in  
FI tag io  
FI tag ip  
FI tag iq  
FI tag ir  
FI tag is  
FI tag it  
FI tag iu  
FI tag iv  
FI tag iw  
FI tag ix  
FI tag iy  
FI tag iz  
FI tag ja  
FI tag jb  
FI tag jc  
FI tag jd  
FI tag je  
FI tag jf  
FI tag jg  
FI tag jh  
FI tag ji  
FI tag jj  
FI tag jk  
FI tag jl  
FI tag jm  
FI tag jn  
FI tag jo  
FI tag jp  
FI tag jq  
FI tag jr  
FI tag js  
FI tag jt  
FI tag ju  
FI tag jv  
FI tag jw  
FI tag jx  
FI tag jy  
FI tag jz  
FI tag ka  
FI tag kb  
FI tag kc  
FI tag kd  
FI tag ke  
FI tag kf  
FI tag kg  
FI tag kh  
FI tag ki  
FI tag kj  
FI tag kk  
FI tag kl  
FI tag km  
FI tag kn  
FI tag ko  
FI tag kp  
FI tag kq  
FI tag kr  
FI tag ks  
FI tag kt  
FI tag ku  
FI tag kv  
FI tag kw  
FI tag kx  
FI tag ky  
FI tag kz  
FI tag la  
FI tag lb  
FI tag lc  
FI tag ld  
FI tag le  
FI tag lf  
FI tag lg  
FI tag lh  
FI tag li  
FI tag lj  
FI tag lk  
FI tag ll  
FI tag lm  
FI tag ln  
FI tag lo  
FI tag lp  
FI tag lq  
FI tag lr  
FI tag ls  
FI tag lt  
FI tag lu  
FI tag lv  
FI tag lw  
FI tag lx  
FI tag ly  
FI tag lz  
FI tag ma  
FI tag mb  
FI tag mc  
FI tag md  
FI tag me  
FI tag mf  
FI tag mg  
FI tag mh  
FI tag mi  
FI tag mj  
FI tag mk  
FI tag ml  
FI tag mn  
FI tag mo  
FI tag mp  
FI tag mq  
FI tag mr  
FI tag ms  
FI tag mt  
FI tag mu  
FI tag mv  
FI tag mw  
FI tag mx  
FI tag my  
FI tag mz  
FI tag na  
FI tag nb  
FI tag nc  
FI tag nd  
FI tag ne  
FI tag nf  
FI tag ng  
FI tag nh  
FI tag ni  
FI tag nj  
FI tag nk  
FI tag nl  
FI tag nm  
FI tag no  
FI tag np  
FI tag nq  
FI tag nr  
FI tag ns  
FI tag nt  
FI tag nu  
FI tag nv  
FI tag nw  
FI tag ox  
FI tag oy  
FI tag oz  
FI tag pa  
FI tag pb  
FI tag pc  
FI tag pd  
FI tag pe  
FI tag pf  
FI tag pg  
FI tag ph  
FI tag pi  
FI tag pj  
FI tag pk  
FI tag pl  
FI tag pm  
FI tag pn  
FI tag po  
FI tag pp  
FI tag pq  
FI tag pr  
FI tag ps  
FI tag pt  
FI tag pu  
FI tag pv  
FI tag pw  
FI tag px  
FI tag py  
FI tag pz  
FI tag qa  
FI tag qb  
FI tag qc  
FI tag qd  
FI tag qe  
FI tag qf  
FI tag qg  
FI tag qh  
FI tag qi  
FI tag qj  
FI tag qk  
FI tag ql  
FI tag qm  
FI tag qn  
FI tag qo  
FI tag qp  
FI tag qq  
FI tag qr  
FI tag qs  
FI tag qt  
FI tag qu  
FI tag qv  
FI tag qw  
FI tag qx  
FI tag qy  
FI tag qz  
FI tag ra  
FI tag rb  
FI tag rc  
FI tag rd  
FI tag re  
FI tag rf  
FI tag rg  
FI tag rh  
FI tag ri  
FI tag rj  
FI tag rk  
FI tag rl  
FI tag rm  
FI tag rn  
FI tag ro  
FI tag rp  
FI tag rq  
FI tag rr  
FI tag rs  
FI tag rt  
FI tag ru  
FI tag rv  
FI tag rw  
FI tag rx  
FI tag ry  
FI tag rz  
FI tag sa  
FI tag sb  
FI tag sc  
FI tag sd  
FI tag se  
FI tag sf  
FI tag sg  
FI tag sh  
FI tag si  
FI tag sj  
FI tag sk  
FI tag sl  
FI tag sm  
FI tag sn  
FI tag so  
FI tag sp  
FI tag sq  
FI tag sr  
FI tag ss  
FI tag st  
FI tag su  
FI tag sv  
FI tag sw  
FI tag sx  
FI tag sy  
FI tag sz  
FI tag ta  
FI tag tb  
FI tag tc  
FI tag td  
FI tag te  
FI tag tf  
FI tag tg  
FI tag th  
FI tag ti  
FI tag tj  
FI tag tk  
FI tag tl  
FI tag tm  
FI tag tn  
FI tag to  
FI tag tp  
FI tag tq  
FI tag tr  
FI tag ts  
FI tag tt  
FI tag tu  
FI tag tv  
FI tag tw  
FI tag tx  
FI tag ty  
FI tag tz  
FI tag ua  
FI tag ub  
FI tag uc  
FI tag ud  
FI tag ue  
FI tag uf  
FI tag ug  
FI tag uh  
FI tag ui  
FI tag uj  
FI tag uk  
FI tag ul  
FI tag um  
FI tag un  
FI tag uo  
FI tag up  
FI tag uq  
FI tag ur  
FI tag us  
FI tag ut  
FI tag uu  
FI tag uv  
FI tag vw  
FI tag vx  
FI tag vy  
FI tag vz  
FI tag wa  
FI tag wb  
FI tag wc  
FI tag wd  
FI tag we  
FI tag wf  
FI tag wg  
FI tag wh  
FI tag wi  
FI tag wj  
FI tag wk  
FI tag wl  
FI tag wm  
FI tag wn  
FI tag wo  
FI tag wp  
FI tag wq  
FI tag wr  
FI tag ws  
FI tag wt  
FI tag wu  
FI tag wv  
FI tag wx  
FI tag wy  
FI tag wz  
FI tag xa  
FI tag xb  
FI tag xc  
FI tag xd  
FI tag xe  
FI tag xf  
FI tag xg  
FI tag xh  
FI tag xi  
FI tag xj  
FI tag xk  
FI tag xl  
FI tag xm  
FI tag xn  
FI tag xo  
FI tag xp  
FI tag xq  
FI tag xr  
FI tag xs  
FI tag xt  
FI tag xu  
FI tag xv  
FI tag xw  
FI tag xx  
FI tag xy  
FI tag yz  
FI tag za  
FI tag zb  
FI tag zc  
FI tag zd  
FI tag ze  
FI tag zf  
FI tag zg  
FI tag zh  
FI tag zi  
FI tag zj  
FI tag zk  
FI tag zl  
FI tag zm  
FI tag zn  
FI tag zo  
FI tag zp  
FI tag zq  
FI tag zr  
FI tag zs  
FI tag zt  
FI tag zu  
FI tag zv  
FI tag zw  
FI tag zx  
FI tag zy  
FI tag zz

misc\_feature 182..190  
/tag= e  
/note= Potential N-linked glycosylation site (AA  
182..190)  
misc\_feature 312..343  
/tag= f  
/note= Potential N-linked glycosylation site (AA  
312..343)  
misc\_feature 442..442  
/tag= g  
/note= Potential N-linked glycosylation site (AA  
442..442)  
misc\_feature 105..107  
/tag= h  
/note= Potential N-linked glycosylation site (AA  
105..107)  
misc\_feature 761..769  
/tag= i  
/note= Potential N-linked glycosylation site (AA  
761..769)  
misc\_feature 214..216  
/tag= j  
/note= Potential N-linked glycosylation site (AA  
214..216)  
misc\_feature 971..979  
/tag= k  
/note= Potential N-linked glycosylation site (AA  
971..979)  
misc\_feature 1280..1288  
/tag= l  
/note= Potential N-linked glycosylation site (AA  
1280..1288)  
misc\_feature 1343..1351  
/tag= m  
/note= Potential N-linked glycosylation site (AA  
1343..1351)  
misc\_feature 408..410  
/tag= n  
/note= Potential N-linked glycosylation site (AA  
408..410)  
misc\_feature 1850..1856  
/tag= o  
/note= Potential N-linked glycosylation site (AA  
1850..1856)  
misc\_feature 577..579  
/tag= p  
/note= Potential N-linked glycosylation site (AA  
577..579)  
misc\_feature 1895..1903  
/tag= q  
/note= Potential N-linked glycosylation site (AA  
1895..1903)  
misc\_feature 593..594  
/tag= r  
/note= Potential N-linked glycosylation site (AA  
593..594)  
misc\_feature 1949..1957  
/tag= s  
/note= Potential N-linked glycosylation site (AA  
1949..1957)  
misc\_feature 2240..2251  
/tag= t  
/note= Potential N-linked glycosylation site (AA  
2240..2251)  
misc\_feature 7012..7103  
/tag= u  
/note= Putative proreceptor processing site (AA  
7012..7103)  
/tag= v  
/note= Beta subunit (AA 711-1337)\*

misc\_feature 2270..2278  
/tag= x  
/note= Potential N-linked glycosylation site (AA  
2270..2278)  
misc\_feature 2397..2405  
/tag= y  
/note= Potential N-linked glycosylation site (AA  
2397..2405)  
misc\_feature 2321..2329  
/tag= z  
/note= Potential N-linked glycosylation site (AA  
2321..2329)  
misc\_feature 734..736  
/tag= aa  
/note= Potential N-linked glycosylation site (AA  
734..736)  
misc\_feature 2729..2737  
/tag= ab  
/note= Potential N-linked glycosylation site (AA  
2729..2737)  
misc\_feature 2768..2776  
/tag= ac  
/note= Potential N-linked glycosylation site (AA  
2768..2776)  
misc\_feature 2837..2808  
/tag= ad  
/note= Potential N-linked glycosylation site (AA  
2837..2808)  
misc\_feature 2918..2926  
/tag= ae  
/note= Transmembrane region (AA 906-929)\*  
misc\_feature 3047..3049  
/tag= af  
/note= Potential N-linked glycosylation site (A\*  
3047..3049)  
misc\_feature 3062..3064  
/tag= ag  
/note= Potential ATP binding site (AA 976)\*  
misc\_feature 3128..3130  
/tag= ah  
/note= Potential ATP binding site (AA 978)\*  
misc\_feature 3128..3130  
/tag= ai  
/note= Potential ATP binding site (AA 981)\*  
misc\_feature 3128..3130  
/tag= aj  
/note= Potential ATP binding site (AA 1003)\*  
US0601515 A1.  
PS 25-JUL-1995; AU0410.  
PP 06-JUL-1995; AU0410.  
PA (ROYA-) ROYAL CHILDRENS HOSPITAL RES FOUND.  
PI Wright GA, Wright CJ;  
DE Wnt3, a member of the Wnt family of proteins, is a secreted  
PT - e.g. ichthyosis, pityriasis and psoriasis, using nucleic  
PT acid which modulates growth factor interaction with its receptor  
PT disclosure: Page 80-82; 116pp; English.  
CC This disclosure describes a method for the treatment of cancer by  
CC the administration of oligonucleotides complementary to the  
CC complete sequence of insulin like growth factor (IGF-1). Keratinocyte growth factor;



```
CC TGF-Alpha: TNO-alpha; Interleukin-1 (IL-1); IL-4, IL-6, IL-8 and bFGF. The effects of proliferative and/or inflammatory skin disorder in a mammal can be ameliorated by administering a nucleic acid or its chemical analogue to inhibit or reduce cell proliferation and angiogenesis. This may include antibodies or otherwise interferes with tcr-i's interaction with its receptor. It is preferable an antisense molecule which reduces expression of the gene encoding tcr-i, an tcr-i receptor or an igf binding protein.
```

```
CC CC CC CC CC CC CC CC  
CG Sequence 9980 BP are numbered G3447, 3326-47, 1082 Z;  
GC Sequence 4980 BP are numbered C3248, 3326-47, 1082 Z;  
ATC i.3%; Score 14; DB 24; Length 4989;
```

```
D Query Match  
D Local Similarity 100.0%; Pred No. 3.47e+02; Oligos 2;  
Methase 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

```
Cb      188 gctgtgggaactgc 201  
Cc      111 cccTGGGAAACACTOC 98  
  
        |||||  
Cb          188 gctgtgggaactgc 201  
Cc          111 cccTGGGAAACACTOC 98
```

```
RESULT 519  
AD R94933 standard; DNA; 5042 BP.  
AD R94933 standard; RNA; 5042 BP.  
DT 20-APR-1998 (first entry)  
DE Porcine heart-fatty acid binding protein gene.  
KW porcin; pig; swine;heart-fatty acid binding protein; H-FABP;  
KW quantitative trait locus; transgenic animal;  
KW fat reduction; ss  
OS Sus scrofa.  
IN Ensembl  
FH Feature  
FT Exon  
FT Exon /tag= a  
FT Exon /number= 1  
FT Exon 2450..2628  
FT Intron /tag= b  
FT Exon /number= 2  
FT Exon 4135..4236  
FT Exon /tag= c  
FT Exon /number= 3  
FT Exon 4768..4821  
FT Exon /tag= d  
FT Exon /number= 4  
FT Exon /start= 4721...1325  
FT misc_feature /note= "polymorphic HindI site"  
FT misc_feature /tag= e  
FT misc_feature /note= "Polymorphic microsatellite sequence (poly-T)".  
FT misc_feature 3128..3331
```

Tue Nov 17 08:55:26 1998  
 US-08-887-977-9.ing

```

CC using samples large numbers of Pst can be genotyped rapidly. The
CC methods can be used to identify differences between strains associated
CC with improvements in production traits, and in marker
CC associated identification/selection of pigs. The H-PASP gene can also
CC be used to generate transgenic animals encoding desirable alleles
CC of H-PASP.
CC Specific allelic proteins or peptides may also be produced from the
CC gene. Such peptides, or antibodies directed against them, can be
CC used to influence production traits in live animals or cell
CC cultures.
CC Desired forms of allelic proteins/fragments can be recombinantly
CC expressed in bacteria, yeast, or mammalian cells.
CC desirable in pig breeding because of interest in lean meat, but
CC previous efforts to reduce backfat deposits have also decreased
CC muscle specific candidate loci for quantitative trait locus mapping.
CC The H-PASP gene is a
CC potential candidate for genetic selection of improved lean meat
CC in pigs. Variation in this gene is responsible for variation in
CC IMF % and backfat thickness, and will affect other variation in
CC fat production, average daily weight gain and feed efficiency, since
CC IMF in relation to 3 polymorphisms in H-PASP, Mapt, Haelll and
CC energy from other purposes, e.g. analysis of body weight (BW) and
CC IMF in relation to 3 polymorphisms in H-PASP, Mapt, Haelll and
CC short by Statistical Analysis System (SAS) statistical programme
CC which was used to analyse the data. There is a significant difference in BW, and a distinct but not
CC significant difference in IMF %.
CC Sequence 5042 BP: 1225 A; 1243 C; 1291 G; 1277 T;
CC SQ
Query Match Score 14.38, Score 14. DB 38, Length 5042;
Best Local Similarity 100.0%; Pos Mod 3.4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2713 slttgttagaat 776
Cp 768 ATTGTATAGGAAT 776

RESULT 520
ID Q79074 standard; DNA: 5169 BP.
AC Q79074:
PF C7-AUG-1995; (first entry)
KW Pl gene: cytochrome; mycoplasma; hybridization; probe. ds
NC Mycobacterium pneumoniae strain M139
FR Key Location/Qualifiers
FF 229..4864
CD /tag-a
PR US5369005-A.
PD 29-NOV-1994.
PR 09-JAN-1997; US-004767.
PR 19-NOV-1987; US-118967.
PR 27-JULY-1990; US-55886B.

```

Tue Nov 17 08:55:26 1998  
US-08-887-977-9.jpg

FT		/note= 'g	
FT		/note= 'polymorphic HaeIII site'	
FT	misc_feature	5035..3904	
FT		/tag= h	
FT		/note= 'non-depicted 3.5 kb intron sequence'	
FT	misc_feature	4012..4013	
FT		/tag= i	
FT		/note= 'non-depicted 0.4 kb intron sequence'	
FT	misc_feature	4567..4568	
FT		/tag= k	
FT	misc_signal	1025..1439	
FT		/tag= l	
FT		/note= '11 nucleotide element'	
FT	misc_signal	1489..1501	
FT		/tag= m	
FT		/note= '13 nucleotide element'	
FT	TATA_signal	1546..1551	
FT		/tag= n	
FT		/note= 'putative TATA-box'	
FT	polya_signal	5025..5031	
FT		/tag= o	
FT		/note= 'putative polyadenylation signal'	
FT		W07158708.42.	
FT	PP	27-MAR-1997; NM157.	
FT	PP	28-MAR-1996; EP-200855.	
FT	PP	(DALLS) DALLAND BV.	
FT	PA	(ENVE) NOORD NEDERLANDS VARENSTAMBOEK BV.	
FT	PA	(PROV.) PROVA BV.	
FT	PA	(STAN-) STAMBOEK ZUID BV.	
FT	PA	W07158707.469563/45.	
FT	PP	W07158707.469563/45.	
FT	DR	W07158707.469563/45.	
FT	FT	Pig heart fatty acid-binding protein gene - used to identify	
FT	FT	polymorphisms associated with production traits, e.g. body weight,	
FT	FT	Glavin 1; Fig 1; 41pp; Encoder.	
FT	PS	The present sequence encodes porcine heart-fatty acid binding	
FT	CC	protein (h-FABP). The h-FABP gene can be used to localise, identify	
FT	CC	and/or to identify polymorphisms associated with production traits,	
FT	CC	especially where these are associated with production traits.	
FT	CC	Alleles of the porcine h-FABP gene can be marked, allowing them to	
FT	CC	be distinguished, preferably by detection of specific restriction	
FT	CC	sites, e.g. MspI, HaeIII or HinfI. The h-FABP gene can also be used	
FT	CC	to identify polymorphisms associated with production traits, by	
FT	CC	loci in samples by amplification of specific genomic fragments. By	

Tue Nov 17 08:55:26 1998  
115-08-887-977+9 vna

[illegible]



The *S. wendlandensis* fosfomycin biosynthesis related gene T06001, which encodes the methyltransferase (MT) and oxyphosphorylation (OP) enzymes R53187 and R53148 respectively, is closely related to the fosfomycin biosynthesis clusters of the MT and OP genes of *S. wendlandensis* (GenBank accession numbers: AF040000 and AF040001). Its plasmid pF0623 can be used to prepare the *S. lividans* variant *S. lividans* pF0623 (FERM P-1423). The prep. variant can be used to prepare fosfomycin by culture in a nutrient medium by conventional methods. Sequence: 5169 bp; 890 A; 1752 G; 75% T.

05-JAN-1991, US-465792.  
(TEZA.) UNIV TEXAS  
Baseman JB, Dello SF, Su CJ;  
WPI: 94-04289/05.  
P-PSDB: RYV11  
The following fragments of Mycoplasma pneumoniae cyradshecin -  
P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13, P14, P15, P16, P17, P18, P19, P20, P21, P22, P23, P24, P25, P26, P27, P28, P29, P30, P31, P32, P33, P34, P35, P36, P37, P38, P39, P40, P41, P42, P43, P44, P45, P46, P47, P48, P49, P50, P51, P52, P53, P54, P55, P56, P57, P58, P59, P60, P61, P62, P63, P64, P65, P66, P67, P68, P69, P70, P71, P72, P73, P74, P75, P76, P77, P78, P79, P80, P81, P82, P83, P84, P85, P86, P87, P88, P89, P90, P91, P92, P93, P94, P95, P96, P97, P98, P99, P100, P101, P102, P103, P104, P105, P106, P107, P108, P109, P110, P111, P112, P113, P114, P115, P116, P117, P118, P119, P120, P121, P122, P123, P124, P125, P126, P127, P128, P129, P130, P131, P132, P133, P134, P135, P136, P137, P138, P139, P140, P141, P142, P143, P144, P145, P146, P147, P148, P149, P150, P151, P152, P153, P154, P155, P156, P157, P158, P159, P160, P161, P162, P163, P164, P165, P166, P167, P168, P169, P170, P171, P172, P173, P174, P175, P176, P177, P178, P179, P180, P181, P182, P183, P184, P185, P186, P187, P188, P189, P190, P191, P192, P193, P194, P195, P196, P197, P198, P199, P200, P201, P202, P203, P204, P205, P206, P207, P208, P209, P210, P211, P212, P213, P214, P215, P216, P217, P218, P219, P220, P221, P222, P223, P224, P225, P226, P227, P228, P229, P230, P231, P232, P233, P234, P235, P236, P237, P238, P239, P240, P241, P242, P243, P244, P245, P246, P247, P248, P249, P250, P251, P252, P253, P254, P255, P256, P257, P258, P259, P260, P261, P262, P263, P264, P265, P266, P267, P268, P269, P270, P271, P272, P273, P274, P275, P276, P277, P278, P279, P280, P281, P282, P283, P284, P285, P286, P287, P288, P289, P290, P291, P292, P293, P294, P295, P296, P297, P298, P299, P300, P301, P302, P303, P304, P305, P306, P307, P308, P309, P310, P311, P312, P313, P314, P315, P316, P317, P318, P319, P320, P321, P322, P323, P324, P325, P326, P327, P328, P329, P330, P331, P332, P333, P334, P335, P336, P337, P338, P339, P340, P341, P342, P343, P344, P345, P346, P347, P348, P349, P350, P351, P352, P353, P354, P355, P356, P357, P358, P359, P360, P361, P362, P363, P364, P365, P366, P367, P368, P369, P370, P371, P372, P373, P374, P375, P376, P377, P378, P379, P380, P381, P382, P383, P384, P385, P386, P387, P388, P389, P390, P391, P392, P393, P394, P395, P396, P397, P398, P399, P400, P401, P402, P403, P404, P405, P406, P407, P408, P409, P410, P411, P412, P413, P414, P415, P416, P417, P418, P419, P420, P421, P422, P423, P424, P425, P426, P427, P428, P429, P430, P431, P432, P433, P434, P435, P436, P437, P438, P439, P440, P441, P442, P443, P444, P445, P446, P447, P448, P449, P450, P451, P452, P453, P454, P455, P456, P457, P458, P459, P460, P461, P462, P463, P464, P465, P466, P467, P468, P469, P470, P471, P472, P473, P474, P475, P476, P477, P478, P479, P480, P481, P482, P483, P484, P485, P486, P487, P488, P489, P490, P491, P492, P493, P494, P495, P496, P497, P498, P499, P500, P501, P502, P503, P504, P505, P506, P507, P508, P509, P510, P511, P512, P513, P514, P515, P516, P517, P518, P519, P520, P521, P522, P523, P524, P525, P526, P527, P528, P529, P530, P531, P532, P533, P534, P535, P536, P537, P538, P539, P540, P541, P542, P543, P544, P545, P546, P547, P548, P549, P550, P551, P552, P553, P554, P555, P556, P557, P558, P559, P560, P561, P562, P563, P564, P565, P566, P567, P568, P569, P570, P571, P572, P573, P574, P575, P576, P577, P578, P579, P580, P581, P582, P583, P584, P585, P586, P587, P588, P589, P590, P591, P592, P593, P594, P595, P596, P597, P598, P599, P600, P601, P602, P603, P604, P605, P606, P607, P608, P609, P610, P611, P612, P613, P614, P615, P616, P617, P618, P619, P620, P621, P622, P623, P624, P625, P626, P627, P628, P629, P630, P631, P632, P633, P634, P635, P636, P637, P638, P639, P640, P641, P642, P643, P644, P645, P646, P647, P648, P649, P650, P651, P652, P653, P654, P655, P656, P657, P658, P659, P660, P661, P662, P663, P664, P665, P666, P667, P668, P669, P670, P671, P672, P673, P674, P675, P676, P677, P678, P679, P680, P681, P682, P683, P684, P685, P686, P687, P688, P689, P690, P691, P692, P693, P694, P695, P696, P697, P698, P699, P700, P701, P702, P703, P704, P705, P706, P707, P708, P709, P710, P711, P712, P713, P714, P715, P716, P717, P718, P719, P720, P721, P722, P723, P724, P725, P726, P727, P728, P729, P730, P731, P732, P733, P734, P735, P736, P737, P738, P739, P740, P741, P742, P743, P744, P745, P746, P747, P748, P749, P750, P751, P752, P753, P754, P755, P756, P757, P758, P759, P760, P761, P762, P763, P764, P765, P766, P767, P768, P769, P770, P771, P772, P773, P774, P775, P776, P777, P778, P779, P780, P781, P782, P783, P784, P785, P786, P787, P788, P789, P790, P791, P792, P793, P794, P795, P796, P797, P798, P799, P800, P801, P802, P803, P804, P805, P806, P807, P808, P809, P810, P811, P812, P813, P814, P815, P816, P817, P818, P819, P820, P821, P822

[illegible]

CC	genitalium.
CC	Seq Q127518-20.
CC	Sequence 5169 BP;
SQ	1323 A; 1454 C; 1279 G; 1113 T;
	Query Match      Score 14; DB 2; Length 5169;
	Best Local Similarity 100.0%; Pred. No. 3.47e+02;
	Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dd	1675 gcacgtcattg 1688       
Cp	245 GCATGCCCATGTT 232
	RESULT 524
ID	Q12752 standard; DWA; 5186 BP.
AC	Q12752:
DT	04-OCT-1991 (first entry)
DZ	Deleted in Colorectal Carcinomas gene.
DE	Diagnosis; antibodies, tumorigenesis, neoplasia; se.
OS	Homo sapiens.
PH	Key
FO	Location/Qualifiers



Query Match 1.38; Score 14; DB 2; Length 5186;  
Best Local Similarity 100.0%; Pred. No. 3,478+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1387 ttgtttcttgcgt 1400  
QY 755 ttgtttcttgcgt 768

RESULT 525  
ID T00333 standard; DNA: 5198 BP.  
DE TM gene.  
KW TM: timeless protein; nuclear translocation protein; circadian rhythm;  
KW sleep/wake cycle; light sensitive; environmental cycle; nuclear protein;  
KW narcolepsy; PER; jet lag; MYP; therapy; ss.  
OS Drosophila melanogaster.  
FH Key Location/Qualifiers  
FT exon /tag= 555  
FT cds /number= 1  
FT /tag= b  
FT /tag= b unsplined TIM  
FT misc\_feature 2387..2450  
FT /tag= c  
FT intron /note= "deleted in tim 01 variant"  
FT /tag= 3556..3792  
FT /tag= 3793..4646  
FT /tag= e  
FT /tag= e  
FT /number= 1  
FT /number= 2  
PD 26-SEP-1996  
PP W09629406-A2  
PP 20-MAR-1996; U03810.  
PP 20-MAR-1995; U5-408518.  
PP 03-NOV-1995; U5-512154.  
PA (UTP) UNIV PENNSYLVANIA.  
PA (UTP) UNIV ROCKEFELLER.  
PI Myers MG, Price JL, Sehgal A, Voshell LB, Young MW;  
P-PDB: W14057.  
PT Nuclear translocation protein which binds to protein involved in  
PT circadian rhythm - used to develop prods. for diagnosis, prevention  
PT or treatment of disorders associated with circadian rhythms, e.g.  
PT jetlag and narcolepsy

Query Match 1.38; Score 14; DB 2; Length 5186;  
Best Local Similarity 100.0%; Pred. No. 3,478+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1387 ttgtttcttgcgt 1400  
QY 755 ttgtttcttgcgt 768

RESULT 525  
ID T00333 standard; DNA: 5198 BP.  
DE TM gene.  
KW TM: timeless protein; nuclear translocation protein; circadian rhythm;  
KW sleep/wake cycle; light sensitive; environmental cycle; nuclear protein;  
KW narcolepsy; PER; jet lag; MYP; therapy; ss.  
OS Drosophila melanogaster.  
FH Key Location/Qualifiers  
FT exon /tag= 555  
FT cds /number= 1  
FT /tag= b  
FT /tag= b unsplined TIM  
FT misc\_feature 2387..2450  
FT /tag= c  
FT intron /note= "deleted in tim 01 variant"  
FT /tag= 3556..3792  
FT /tag= 3793..4646  
FT /tag= e  
FT /tag= e  
FT /number= 1  
FT /number= 2  
PD 26-SEP-1996  
PP W09629406-A2  
PP 20-MAR-1996; U03810.  
PP 20-MAR-1995; U5-408518.  
PP 03-NOV-1995; U5-512154.  
PA (UTP) UNIV PENNSYLVANIA.  
PA (UTP) UNIV ROCKEFELLER.  
PI Myers MG, Price JL, Sehgal A, Voshell LB, Young MW;  
P-PDB: W14057.  
PT Nuclear translocation protein which binds to protein involved in  
PT circadian rhythm - used to develop prods. for diagnosis, prevention  
PT or treatment of disorders associated with circadian rhythms, e.g.  
PT jetlag and narcolepsy

Query Match 1.38; Score 14; DB 14; Length 5248;  
Best Local Similarity 100.0%; Pred. No. 3,478+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 184 ttggggcttgcgt 197  
QY 616 ttggggcttgcgt 629

RESULT 528  
ID T00189 standard; cDNA: 5285 BP.  
DE TM gene.  
KW TM: timeless protein; nuclear translocation protein; circadian rhythm;  
KW sleep/wake cycle; light sensitive; environmental cycle; nuclear protein;  
KW narcolepsy; PER; jet lag; MYP; therapy; ss.  
OS Drosophila melanogaster.  
FH Key Location/Qualifiers  
FT exon /tag= 555  
FT cds /number= 1  
FT /tag= b  
FT /tag= b unsplined TIM  
FT misc\_feature 2387..2450  
FT /tag= c  
FT intron /note= "deleted in tim 01 variant"  
FT /tag= 3556..3792  
FT /tag= 3793..4646  
FT /tag= e  
FT /tag= e  
FT /number= 1  
FT /number= 2  
PD 26-SEP-1996  
PP W09629406-A2  
PP 20-MAR-1996; U03810.  
PP 20-MAR-1995; U5-408518.  
PP 03-NOV-1995; U5-512154.  
PA (UTP) UNIV PENNSYLVANIA.  
PA (UTP) UNIV ROCKEFELLER.  
PI Myers MG, Price JL, Sehgal A, Voshell LB, Young MW;  
P-PDB: W14057.  
PT Nuclear translocation protein which binds to protein involved in  
PT circadian rhythm - used to develop prods. for diagnosis, prevention  
PT or treatment of disorders associated with circadian rhythms, e.g.  
PT jetlag and narcolepsy

Query Match 1.38; Score 14; DB 13; Length 5238;  
Best Local Similarity 100.0%; Pred. No. 3,478+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3426 ttggggcttgcgt 3439  
Cp 1019 ttggggcttgcgt 1006

RESULT 527  
ID T00189 standard; DNA: 5248 BP.  
DE TM gene.  
KW TM: timeless protein; nuclear translocation protein; circadian rhythm;  
KW sleep/wake cycle; light sensitive; environmental cycle; nuclear protein;  
KW narcolepsy; PER; jet lag; MYP; therapy; ss.  
OS Drosophila melanogaster.  
FH Key Location/Qualifiers  
FT exon /tag= 555  
FT cds /number= 1  
FT /tag= b  
FT /tag= b unsplined TIM  
FT misc\_feature 2387..2450  
FT /tag= c  
FT intron /note= "deleted in tim 01 variant"  
FT /tag= 3556..3792  
FT /tag= 3793..4646  
FT /tag= e  
FT /tag= e  
FT /number= 1  
FT /number= 2  
PD 26-SEP-1996  
PP W09629406-A2  
PP 20-MAR-1996; U03810.  
PP 20-MAR-1995; U5-408518.  
PP 03-NOV-1995; U5-512154.  
PA (UTP) UNIV PENNSYLVANIA.  
PA (UTP) UNIV ROCKEFELLER.  
PI Myers MG, Price JL, Sehgal A, Voshell LB, Young MW;  
P-PDB: W14057.  
PT Nuclear translocation protein which binds to protein involved in  
PT circadian rhythm - used to develop prods. for diagnosis, prevention  
PT or treatment of disorders associated with circadian rhythms, e.g.  
PT jetlag and narcolepsy











Agaricillus oryzae strain IPO 4178.  
Key:        C=conserved  
CD5

2382..4951  
          /ctg=a  
          /ctg=b  
          /ctg=c  
          /ctg=d  
          /ctg=e  
          /ctg=f  
          /ctg=g  
          /ctg=h  
          /ctg=i  
          /ctg=j  
          /ctg=k  
          /ctg=l  
          /ctg=m  
          /ctg=n  
          /ctg=o  
          /ctg=p  
          /ctg=q  
          /ctg=r  
          /ctg=s  
          /ctg=t  
          /ctg=u  
          /ctg=v  
          /ctg=w  
          /ctg=x  
          /ctg=y  
          /ctg=z

exon  
intron  
exon

W09723705-AL  
26-JUN-1997  
16-DEC-1986; DK0528.  
PR 15-DEC-1995; DK-001428.  
PR 15-DEC-1995; DK-001428.  
PA Christensen T, Møller A S,  
WPI: 97-341696/J1.  
P-PS9B; W11630.

new fungus lacking functional Area and extracellular protease  
proteins, also new protease(s) from Agaricillus oryzae  
Example 3: Page 40-45; 80pp: English.  
This DNA sequence includes a coding region for the Area regulator  
gene. The invention relates to a new fungus in which: (a) the Area  
gene was cloned by cross-hybridization with the Area gene  
gene has been modified by recombinant DNA methods so that it cannot  
express a functional Area activator; and (2) the extracellular  
inactivated so that functional enzyme is not produced. The  
fungi have complete or partial deletions of the appropriate genes,  
or expression sequences that regulate the genes. Alternatively,  
the genes are replaced by other genes. The new fungi are used for the recombinant production  
of industrial or therapeutic peptides and proteins, particularly  
enzymes (e.g. proteases, lipase, cutinase, cellulase or chymotrypsin),  
including growth hormone, glucagon, somatostatin, interferon, cytokines,  
tissue plasminogen activator, erythropoietin or thrombopoietin.  
The new fungi do not produce protease and so provide higher yields  
of recombinant protein because of reduced proteolytic degradation.  
Sequence 3843 BP: 1199 A: 1618 C: 1348 G: 1478 T:

Query Match  
Best Local Similarity 100.0%; Fred No. 3-47e+02;  
Matches 14; Conservative 0 Mismatches 0 Indels 0 Gaps

0-000000

```

DR  P-DBP, R93365.
PT  Kary R21-Binding Transcription factor (RE37 protein) - used
PT  therapeutically to control differentiation and activity of neural
PT  cells, e.g. in cases of brain cancer
PT  Claim 14, Page 30-39; 10pp; English.
PS  Genbank accession number for human RE37 (R93365) or RE1-
CC  binding transcription factors for non-neural tissues, or RE1-
CC  expression of neural proteins in non-neural tissues by silencing
CC  promoter congs. RE1 elements (see also T41407). The clone was
CC  abd. from HeLa cells (499) the one-hybrid cloning of 3 partial
CC  overlapping sequences (T41413, T41413) encoding RE37. It can be used
CC  to produce RE37, or RE37 DNA binding domain (see also R93364), in
CC  transformed host cells, as a probe, and in gene therapy to treat
CC  neurodegenerative diseases involving neural activity; antisense
CC  sequences can be used to treat disorders involving de-differentiated
CC  neural cells.
SQ  Sequence 5648 BP: 1747 A; 1098 C; 1284 G; 1507 T;
      Query Match 1.38; Score 14; DB 24; Length 5648;
      Best Local Similarity 100.0%; Pred Mo. 3.47e+02; Indels 0; Gaps
      Matches 14; Conservative 0; Mismatches 0;
Db  3003 TCTTCAGAAAGTAG 3016
Cp  11111111111111111111
      961 TCTTCAGAAAGTAG 948

RESULT 540
ID  Q13572 standard; DNA; 5706 BP.
AC  Q13572;
DT  29-NOV-1991 (first entry)
KW  RE37 protein; transcription factor; hormone;
OS  Drosophila melanogaster;
FH  Key Location/Qualifiers
FT  cds 631..4404
FT  /translation= a
FT  /tag= b
FT  /misc_difference 3002
FT  /tag= a
FT  /note= "In cDNA clone Dm4925 T=."
FT  /size= 5706
FT  /size= 5395
FT  /tag= b
FT  exon 631..1088
FT  /tag= d
FT  /note= "Exon A0"
FT  /size= 148
FT  /tag= a
FT  /note= "Exon A1"
FT  /size= 1489..1633
FT  exon

```



```

/notes="Exon 2"
exon      1634..1672
          /tag= g
exon      /note="Exon 3"
          /tag= 281
          /tag= 281
exon      /note="Exon 4"
          /tag= 282..4104
          /note="Exon 5"
          /tag= 4902..4907
          /tag= j
          /note="PolyA signal for cDNA clones cdm4927 and
polyA_signal 5674..5679
          /tag= k
          /note="putative polyA signal for long transcripts"
polyA_site /tag= l
          /note="PolyA site for cDNA clones cdm4927 and
          cdm4928"
909113167.4.
05-DEC-1991.
05-DEC-1991.
15-DEC-1991. 001189.
26-FEB-1990. 05-485749.
(STUD ) LELAND STANFORD JR UNIV
(MODEL) HUMAN DRUG METABOLISM MR. SEGRAVAS WA.
(MATCH) 91.281480/18.
P-PSDB: R13791.
cDNA encoding insect steroid receptors - and ligands, for use as
Cloning 2 inducing factors. English.
The DNA sequence encodes the E75A transcription unit which consists
of six exons A0, A1, 2, 3, 4 and 5, of these A0 and A1 are specific
to E75A while 2, 3, 4 and 5 are shared with the E75B transcription
unit. The E75A gene is located on chromosome 2. The E75 gene, it
displays characteristic Drosophila melanogaster codon usage. The
E75A proteins show considerable similarity to members of the steroid
receptor family. Since the putative hormone binding E domain of the
protein contains a high sequence homology to the known E domain of
either a terpenoid inducible protein (Q15174) or a protein that binds
a terpenoid juvenile hormone of a novel Drosophila hormone.
See also Q13573-Q13575.
Sequence 5706 BP. 1570 A. 1591 C. 1325 G. 1220 T.
Query Match 1.3%; Score 14; DB 2; Length 5706;
Best Local Similarity 100.0%; Pred. No. 3.47e-02;
Matches 14: Conservative 0; Mismatches 0; Indels 0; Gaps
0
3777 gggatccaccca 3790
|||||

```

Tue Nov 17 08:55:26 1998

```

      ||||||!!!!!!|||
045 AAGGACCAAGAAG 631

RESULT 542
ID Q1078 standard; cDNA; 5958 BP.
PR 12-APR-1991 (first entry)
EN Encodes human MPOd mediator of inflammation protein.
DE Mediator of Inflammation: cytokine; Hodgkin's lymphoma. MPD-160; as.
KW Homo sapiens.
FT Location/Qualifiers
   cdg     134..441?
          /*tag= "a
           /product= MRP-160
           /note="mediator of inflammation"

BP-A42050-A.
PD 06-FEB-1991.
PR 26-JUN-1990; 810481.
EN GC-BUL-1989; GB-G13414.
DE Olink AG, Tarsney Bruuggen J, Wieseandanger W, Cerletti N.
TI Soryk C, Dewolf-Peters C, Delabie J.
WP1: 91-018913/06.
PR 160 NP. RUS34.
EN Poly(ethylene glyclide medietor or precursor of inflammation -
TI polygonal or monoclonal antibodies to polypeptide treat and
PR diagnose chronic inflammation and hodgkins lymphoma
CLM 14; Page 3; 47pp; English.
CC which was cloned in puen LIZ2 cells wih Thibaut
CC the sequence encoding the molecule whic reacts with MabI C5 (see EP-
CC 162812), by successive screening wit MabI C5 and cDNA inserts from
CC the commercially available cDNA libraries ML105S8 and ML10298. One
CC clone was found. The sequence was determined by direct sequencing. Found
CC to encode MRP-160. The sequene was cloned in pUCMKO to give a hybrid
CC vector which was used to transform E.coli DEB-alpha cells. A plasmid
CC yielding the correct restriction fragments was designated pMRP160.
CC Sequence 5958 BP; 1962 A; 1171 C; 1374 G; 1351 ?;
Query Match 1.3%; Score 15; DB 2; Length 5858;
Best Local Similarity 100.0% Pred.No. 8.37e+01;
Matches 15; Conservative 0; Missmatches 0; Indels 0; Gaps 0;
db 1381 gctgggaactcgtcg 1395
IT 599 GGTGGGAAGTCCTCGA 613

RESULT 543
ID Q86678 standard; cDNA; 6000 BP.

```

**Tue Nov 17 08:55:26 1998**

```

Cp      183 GGTATACACCA 170

RESULT 541
ID M40031 standard; DNA: 5778 BP.
DZ 25-JAN-1992 (first entry)
DE sequence of portion of human beta-nerve growth factor (NGF)
DE chromosome gene which includes an exon.
DE nerve damage; therapy; se.
FH 16-BETA-NGF cDNA
FR 16-beta NGF
Key Location/Qualifiers
misc_feature
    1
/*tag= A 8bb separates this base from the
previous sequenced portion of the human
beta-NGF chromosomal gene in M40031*/
exon       3306..4044
poly_a_signal 5031..5026
poly_a_signal 5261..5266
poly_a_signal 5785..5783
poly_a_signal /*tag= e
EP-121338-A.
PD 10-OCT-1984. 301377
PD 03-MAR-1983. DB-471962
PA (GETH ) GENENTECH INC.
PI Gray AM, Ollrich A.
PI PNAS 84:2408/41.
PD 1983MAY20/41.
DB Human beta-nerve growth factor free from other proteins - obt'd.
PP by recombinant DNA techniques for treating nerve damage
PP Example; Fig 5, 42pp; English. Nerve growth factor (NGF) free from
PP other proteins of human origin. Also claimed are the DNA sequence
PP encoding human beta-NGF operably linked with a DNA sequence capable
PP of affecting its expression in a recombinant host cell; a replicable
PP vector that could be claimed in animal ph-beta-NGF transgene. Using
PP the plasmid. Larger amounts of pure beta-NGF are obtainable than by
PP extra. of natural materials, see e.g. EP-2139.
PP sequence 5778 BP; 16S2 A; 1435 C; 1377 T.
Query Match 1.3%; Score 15; DB 3; Length 5778;
Best Local Similarity 100.0%; Pred. No. 8.17e+01;
Matches 0; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Db 1087 aaagcggccacgaaag 1101

```

Tue Nov 17 08:55:26 1998

```

AC 086478: 17-OCT-1995 (first entry)
DE human PTP-OB
DE human protein tyrosine phosphatase; osteoblast; differentiation;
DE osteoblast; osteopetrosis, bone; cancer; osteosarcoma, ss.
DE Homo sapiens.
PF W059079315.A.
PP 09-SEP-1994.
PP 14-SEP-1993: US-122032.
PP (KERI ) MERCK & CO INC.
PP Rodan GA, Rutledge ST, Schmidt A.
PP P-DBS: R1726/1.
PP Protein tyrosine phosphate protein PTP-OB specifically expressed
PP in bone cells - modulators of which are used to treat, e.g.
PP osteoporosis, 45-67 percent and breast bone loss and cancer.
PP CCR amplification of cDNA derived from human osteosarcoma
PP PCR
CC Sacc-2/10 using primers based on conserved regions of protein
CC tyrosine phosphatases and subsequent screening of a human
CC cDNA library.
CC (Q86473) that encoded a novel human protein, PTP-OB (R1726).
CC Sequence 6000 BP; 1228 A; 1371 C; 1803 G; 987 T;
CC SQ
Query Match 1.38; Score 15; DB 14; Length 6006;
Query Match Similarity 100.38; Pct Ids 8.93; 101.
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 4611 tctgtgcttctgttcg 4625
QY 881 tcttgcgcttctgttcg 895
RESULT 544
AC 785189: standard, cDNA, 6000 BP.
AC 785189:
DB 18-DEC-1997 (first entry)
DE human protein tyrosine phosphatase PTP-OB encoding cDNA.
DE human protein tyrosine phosphatase PTP-OB; Preparation, osteoblast;
DE osteoblast protein, growth, differentiation, brain, human, ss.
DE Homo sapiens.
PF Key Location/Qualifiers
PF cds 21..5758
PF 1000 bp
PF /product= PTP-OB
PF US5658756.A.
PP 19-AUG-1997.
PP 04-SEP-1993: 122032
PP (KERI ) MERCK & CO INC.
PP 14-SEP-1993: US-122032
PP P-DBS: R1726/1.

```



















```

RESULT 557
ID V02042 standard: DNA; 7287 BP.
CD C
DE 08-JUN-1988 (first entry)
DE Plasmid pRC3169 encoding murine interleukin-12.
KW Interleukin-12; IL-12; cytokine; growth factor; mouse;
KW plasmid pRC3169; cancer; tumour; metastasis; gene therapy; ds:
OS Chimeric - Virus musculus.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Bos taurus.
OS Chimeric - Rhesus macaque polyoma virus.
FT Chimeric - Rhesus macaque polyoma virus.
FT Promoter 1..628
/+/tag= a
/+/note= CMV promoter*
LDNA /+/tag= b
CDS 953..1673
/+/tag= c
/+/product= p35 subunit
Intron 1259..1331
/+/tag= d
polyA_site 1797..2024
promoter 2110..2737
/+/tag= f
/+/note= CMV promoter*
LDNA 2738..2919
/+/tag= g
CDS 2983..3980
/+/tag= h
/+/tag= i
polyA_site 4075..4306
/+/tag= j
/+/note= bovine growth hormone polyA site*
W00746363.b1.
PD 11-DIC-1987
PP 04-JUN-1997 U09591.
PP 05-JUN-1986; US-653206.
PA (ACRA); AURAGEN INC.
DA WPI: 98-041898/04
DR P-PSIB: M44004-05.
DR Interleukin-12 gene therapy of tumours - comprises delivering
DR p35 and p40 subunits to target cells in vivo
SI Claim 5: Page 24-30; Sopp: English.
PS

```

```

TUE Nov 17 08:55:26 1998

US-08-887-977-9.XMG

FT mat_peptide 3516..4328 /tag= e
FT /note= 'putative POL protein (partial)'
FT mat_peptide 4322..4748
FT CDS 4738..6725 /tag= g
FT /note= 'putative ENV coding region (partial) as
FT 4752..6725 described in the specification'
FT mat_peptide /tag= h
FT /note= 'ENV protein (partial)'
FT M0971836.A1.
FT 13-DEC-1996: U19680
FT 14-DEC-1995: U572645.
FT (GENE) GEN HOSPITAL CORP.
FT Fishman JA.
FT P1 310694.30
FT P2 310694.30
FT P3 310694.30
FT P4 310694.30
FT P5 310694.30
FT P6 310694.30
FT P7 310694.30
FT P8 310694.30
FT P9 310694.30
FT P10 310694.30
FT P11 310694.30
FT P12 310694.30
FT P13 310694.30
FT P14 310694.30
FT P15 310694.30
FT P16 310694.30
FT P17 310694.30
FT P18 310694.30
FT P19 310694.30
FT P20 310694.30
FT P21 310694.30
FT P22 310694.30
FT P23 310694.30
FT P24 310694.30
FT P25 310694.30
FT P26 310694.30
FT P27 310694.30
FT P28 310694.30
FT P29 310694.30
FT P30 310694.30
FT P31 310694.30
FT P32 310694.30
FT P33 310694.30
FT P34 310694.30
FT P35 310694.30
FT P36 310694.30
FT P37 310694.30
FT P38 310694.30
FT P39 310694.30
FT P40 310694.30
FT P41 310694.30
FT P42 310694.30
FT P43 310694.30
FT P44 310694.30
FT P45 310694.30
FT P46 310694.30
FT P47 310694.30
FT P48 310694.30
FT P49 310694.30
FT P50 310694.30
FT P51 310694.30
FT P52 310694.30
FT P53 310694.30
FT P54 310694.30
FT P55 310694.30
FT P56 310694.30
FT P57 310694.30
FT P58 310694.30
FT P59 310694.30
FT P60 310694.30
FT P61 310694.30
FT P62 310694.30
FT P63 310694.30
FT P64 310694.30
FT P65 310694.30
FT P66 310694.30
FT P67 310694.30
FT P68 310694.30
FT P69 310694.30
FT P70 310694.30
FT P71 310694.30
FT P72 310694.30
FT P73 310694.30
FT P74 310694.30
FT P75 310694.30
FT P76 310694.30
FT P77 310694.30
FT P78 310694.30
FT P79 310694.30
FT P80 310694.30
FT P81 310694.30
FT P82 310694.30
FT P83 310694.30
FT P84 310694.30
FT P85 310694.30
FT P86 310694.30
FT P87 310694.30
FT P88 310694.30
FT P89 310694.30
FT P90 310694.30
FT P91 310694.30
FT P92 310694.30
FT P93 310694.30
FT P94 310694.30
FT P95 310694.30
FT P96 310694.30
FT P97 310694.30
FT P98 310694.30
FT P99 310694.30
FT P100 310694.30
FT P101 310694.30
FT P102 310694.30
FT P103 310694.30
FT P104 310694.30
FT P105 310694.30
FT P106 310694.30
FT P107 310694.30
FT P108 310694.30
FT P109 310694.30
FT P110 310694.30
FT P111 310694.30
FT P112 310694.30
FT P113 310694.30
FT P114 310694.30
FT P115 310694.30
FT P116 310694.30
FT P117 310694.30
FT P118 310694.30
FT P119 310694.30
FT P120 310694.30
FT P121 310694.30
FT P122 310694.30
FT P123 310694.30
FT P124 310694.30
FT P125 310694.30
FT P126 310694.30
FT P127 310694.30
FT P128 310694.30
FT P129 310694.30
FT P130 310694.30
FT P131 310694.30
FT P132 310694.30
FT P133 310694.30
FT P134 310694.30
FT P135 310694.30
FT P136 310694.30
FT P137 310694.30
FT P138 310694.30
FT P139 310694.30
FT P140 310694.30
FT P141 310694.30
FT P142 310694.30
FT P143 310694.30
FT P144 310694.30
FT P145 310694.30
FT P146 310694.30
FT P147 310694.30
FT P148 310694.30
FT P149 310694.30
FT P150 310694.30
FT P151 310694.30
FT P152 310694.30
FT P153 310694.30
FT P154 310694.30
FT P155 310694.30
FT P156 310694.30
FT P157 310694.30
FT P158 310694.30
FT P159 310694.30
FT P160 310694.30
FT P161 310694.30
FT P162 310694.30
FT P163 310694.30
FT P164 310694.30
FT P165 310694.30
FT P166 310694.30
FT P167 310694.30
FT P168 310694.30
FT P169 310694.30
FT P170 310694.30
FT P171 310694.30
FT P172 310694.30
FT P173 310694.30
FT P174 310694.30
FT P175 310694.30
FT P176 310694.30
FT P177 310694.30
FT P178 310694.30
FT P179 310694.30
FT P180 310694.30
FT P181 310694.30
FT P182 310694.30
FT P183 310694.30
FT P184 310694.30
FT P185 310694.30
FT P186 310694.30
FT P187 310694.30
FT P188 310694.30
FT P189 310694.30
FT P190 310694.30
FT P191 310694.30
FT P192 310694.30
FT P193 310694.30
FT P194 310694.30
FT P195 310694.30
FT P196 310694.30
FT P197 310694.30
FT P198 310694.30
FT P199 310694.30
FT P200 310694.30
FT P201 310694.30
FT P202 310694.30
FT P203 310694.30
FT P204 310694.30
FT P205 310694.30
FT P206 310694.30
FT P207 310694.30
FT P208 310694.30
FT P209 310694.30
FT P210 310694.30
FT P211 310694.30
FT P212 310694.30
FT P213 310694.30
FT P214 310694.30
FT P215 310694.30
FT P216 310694.30
FT P217 310694.30
FT P218 310694.30
FT P219 310694.30
FT P220 310694.30
FT P221 310694.30
FT P222 310694.30
FT P223 310694.30
FT P224 310694.30
FT P225 310694.30
FT P226 310694.30
FT P227 310694.30
FT P228 310694.30
FT P229 310694.30
FT P230 310694.30
FT P231 310694.30
FT P232 310694.30
FT P233 310694.30
FT P234 310694.30
FT P235 310694.30
FT P236 310694.30
FT P237 310694.30
FT P238 310694.30
FT P239 310694.30
FT P240 310694.30
FT P241 310694.30
FT P242 310694.30
FT P243 310694.30
FT P244 310694.30
FT P245 310694.30
FT P246 310694.30
FT P247 310694.30
FT P248 310694.30
FT P249 
```

```

CC Plasmid pWRG1169 is a tandem plasmid encoding both subunits, i.e.
CC p35 (see M44004) and p40 (see M44003), of murine interleukin-12
CC (IL-12). Each subunit gene was cloned from a mouse spleen cDNA
CC library and has been placed under the transcriptional control of
CC the SV40 promoter. The backbone of the plasmid is derived from
CC donor/splicing acceptor is provided between each subunit, and
CC its CAG promoter. The backbone of the plasmid is from pUC19.
CC pWRG1193 induces at least twice as much IL-12 expression as
CC pWRG1169 does. pWRG1193 (see W00484) in vivo and in vitro A
CC transcribed construct containing two copies of the SV40 promoter
CC copies of an expressible foreign genetic construct, especially
CC pWRG1169 or pWRG1193, fusing a promoter operative in the
CC mammalian epithelial cells and DNA sequences encoding p35 and p40
CC subunits of IL-12. This construct allows IL-12 expression for treatment of solid
CC tumours. The treatment is effective even when the genetic
CC construct is delivered by a dendritic cell derived from the same tumour; 766 T;
CC Sequence 7487 BP; 1798 A; 1817 C; 180 G; 180 T;
CC
CC Query Match 1.3%; Score 14; DB 39; Length 7287;
CC Best Local Similarity 100.0%; Pred. No. 3,474+02;
CC Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps
CC
Db 1126 ttcctcgtcgctg 1139
QY 889 ttccctccccctgc 902

RESULT 558
ID 774883 standard; cDNA; 7393 BP.
D ID 774883 standard; cDNA; 7393 BP.
DT 09-FEB-1998 (first entry)
DE Porcine retrovirus cDNA (defective).
KW Retrovirus; porcine; GAG protein; POL protein; ENV protein;
KW activated virus; infectious; provirus; organ transplant; donor;
OS activated virus PCB #8.
FE Key Location/Qualifiers
FE Key 598..2172
FE CDS /note= "putative GAG protein"
Ft Ft mat_peptide 598..2163
Ft Ft /tag= b
Ft Ft /note= "413"
Ft Ft /tag= c
Ft Ft /note= "Putative POL coding region (partial) as described in the specification"
Ft Ft mat_peptide 2120..3522
Ft Ft /note= "Putative POL protein (partial)"

```

```

Tue Nov 17 09:15:26 1998

US-08-487-977-9.rmg

RESULT 559
ID M2002 standard; cDNA: 7440 BP.
DT 05-NOV-1992 (first entry)
DT 05-NOV-1992
DE Sequence of a full-length cDNA copy of the poliovirus genome in
D2 Poliovirus; picornavirus; vaccine; antigen; immunogen; ss.
KW Poliovirus; picornavirus; vaccine; antigen; immunogen; ss.
PR Human poliovirus
PH Kansas poliovirus Location/Qualifiers
cds 743..949
    /tag= a
    /product= P4
    /size= 1768
    /tag= b
    /product= VP2
    /size= 1766..2479
    /tag= c
    /product= VP3
    /size= 2480..3385
    /tag= d
    /product= VP1
    /size= 3386..3832
    /tag= e
    /product= 3b
    /size= 3833..4123
    /tag= f
    /product= 5b
    /size= 4124..510
    /tag= g
    /product= X
    /size= 5111..5371
    /tag= h
    /product= 1b
    /size= 5372..5586
    /tag= i
    /product= VPg
    /size= 5587..7375
    /tag= j
    /product= 4(p63)
    /size= 7376..7440
    /tag= k
    /product= 4(p63)
    /size= 7441..7440
    /tag= l
    /product= 4(p63)
    /size= 7441..7440
    /tag= m
    /product= 4(p63)
    /size= 7441..7440
    /tag= n
    /product= 4(p63)
    /size= 7441..7440
    /tag= o
    /product= 4(p63)
    /size= 7441..7440
    /tag= p
    /product= 4(p63)
    /size= 7441..7440
    /tag= q
    /product= 4(p63)
    /size= 7441..7440
    /tag= r
    /product= 4(p63)
    /size= 7441..7440
    /tag= s
    /product= 4(p63)
    /size= 7441..7440
    /tag= t
    /product= 4(p63)
    /size= 7441..7440
    /tag= u
    /product= 4(p63)
    /size= 7441..7440
    /tag= v
    /product= 4(p63)
    /size= 7441..7440
    /tag= w
    /product= 4(p63)
    /size= 7441..7440
    /tag= x
    /product= 4(p63)
    /size= 7441..7440
    /tag= y
    /product= 4(p63)
    /size= 7441..7440
    /tag= z
    /product= 4(p63)
    /size= 7441..7440
    /tag= aa
    /product= 4(p63)
    /size= 7441..7440
    /tag= ab
    /product= 4(p63)
    /size= 7441..7440
    /tag= ac
    /product= 4(p63)
    /size= 7441..7440
    /tag= ad
    /product= 4(p63)
    /size= 7441..7440
    /tag= ae
    /product= 4(p63)
    /size= 7441..7440
    /tag= af
    /product= 4(p63)
    /size= 7441..7440
    /tag= ag
    /product= 4(p63)
    /size= 7441..7440
    /tag= ah
    /product= 4(p63)
    /size= 7441..7440
    /tag= ai
    /product= 4(p63)
    /size= 7441..7440
    /tag= aj
    /product= 4(p63)
    /size= 7441..7440
    /tag= ak
    /product= 4(p63)
    /size= 7441..7440
    /tag= al
    /product= 4(p63)
    /size= 7441..7440
    /tag= am
    /product= 4(p63)
    /size= 7441..7440
    /tag= an
    /product= 4(p63)
    /size= 7441..7440
    /tag= ao
    /product= 4(p63)
    /size= 7441..7440
    /tag= ap
    /product= 4(p63)
    /size= 7441..7440
    /tag= aq
    /product= 4(p63)
    /size= 7441..7440
    /tag= ar
    /product= 4(p63)
    /size= 7441..7440
    /tag= as
    /product= 4(p63)
    /size= 7441..7440
    /tag= at
    /product= 4(p63)
    /size= 7441..7440
    /tag= au
    /product= 4(p63)
    /size= 7441..7440
    /tag= av
    /product= 4(p63)
    /size= 7441..7440
    /tag= aw
    /product= 4(p63)
    /size= 7441..7440
    /tag= ax
    /product= 4(p63)
    /size= 7441..7440
    /tag= ay
    /product= 4(p63)
    /size= 7441..7440
    /tag= az
    /product= 4(p63)
    /size= 7441..7440
    /tag= ba
    /product= 4(p63)
    /size= 7441..7440
    /tag= bb
    /product= 4(p63)
    /size= 7441..7440
    /tag= bc
    /product= 4(p63)
    /size= 7441..7440
    /tag= bd
    /product= 4(p63)
    /size= 7441..7440
    /tag= be
    /product= 4(p63)
    /size= 7441..7440
    /tag= bf
    /product= 4(p63)
    /size= 7441..7440
    /tag= bg
    /product= 4(p63)
    /size= 7441..7440
    /tag= bh
    /product= 4(p63)
    /size= 7441..7440
    /tag= bi
    /product= 4(p63)
    /size= 7441..7440
    /tag= bj
    /product= 4(p63)
    /size= 7441..7440
    /tag= bk
    /product= 4(p63)
    /size= 7441..7440
    /tag= bl
    /product= 4(p63)
    /size= 7441..7440
    /tag= bm
    /product= 4(p63)
    /size= 7441..7440
    /tag= bn
    /product= 4(p63)
    /size= 7441..7440
    /tag= bo
    /product= 4(p63)
    /size= 7441..7440
    /tag= bp
    /product= 4(p63)
    /size= 7441..7440
    /tag= bq
    /product= 4(p63)
    /size= 7441..7440
    /tag= br
    /product= 4(p63)
    /size= 7441..7440
    /tag= bs
    /product= 4(p63)
    /size= 7441..7440
    /tag= bt
    /product= 4(p63)
    /size= 7441..7440
    /tag= bu
    /product= 4(p63)
    /size= 7441..7440
    /tag= bv
    /product= 4(p63)
    /size= 7441..7440
    /tag= bw
    /product= 4(p63)
    /size= 7441..7440
    /tag= bx
    /product= 4(p63)
    /size= 7441..7440
    /tag= by
    /product= 4(p63)
    /size= 7441..7440
    /tag= bz
    /product= 4(p63)
    /size= 7441..7440
    /tag= ca
    /product= 4(p63)
    /size= 7441..7440
    /tag= cb
    /product= 4(p63)
    /size= 7441..7440
    /tag= cc
    /product= 4(p63)
    /size= 7441..7440
    /tag= cd
    /product= 4(p63)
    /size= 7441..7440
    /tag= ce
    /product= 4(p63)
    /size= 7441..7440
    /tag= cf
    /product= 4(p63)
    /size= 7441..7440
    /tag= cg
    /product= 4(p63)
    /size= 7441..7440
    /tag= ch
    /product= 4(p63)
    /size= 7441..7440
    /tag= ci
    /product= 4(p63)
    /size= 7441..7440
    /tag= cj
    /product= 4(p63)
    /size= 7441..7440
    /tag= ck
    /product= 4(p63)
    /size= 7441..7440
    /tag= cl
    /product= 4(p63)
    /size= 7441..7440
    /tag= cm
    /product= 4(p63)
    /size= 7441..7440
    /tag= cn
    /product= 4(p63)
    /size= 7441..7440
    /tag= co
    /product= 4(p63)
    /size= 7441..7440
    /tag= cp
    /product= 4(p63)
    /size= 7441..7440
    /tag= cq
    /product= 4(p63)
    /size= 7441..7440
    /tag= cr
    /product= 4(p63)
    /size= 7441..7440
    /tag= cs
    /product= 4(p63)
    /size= 7441..7440
    /tag= ct
    /product= 4(p63)
    /size= 7441..7440
    /tag= cu
    /product= 4(p63)
    /size= 7441..7440
    /tag= cv
    /product= 4(p63)
    /size= 7441..7440
    /tag= cw
    /product= 4(p63)
    /size= 7441..7440
    /tag= cx
    /product= 4(p63)
    /size= 7441..7440
    /tag= cy
    /product= 4(p63)
    /size= 7441..7440
    /tag= cz
    /product= 4(p63)
    /size= 7441..7440
    /tag= da
    /product= 4(p63)
    /size= 7441..7440
    /tag= db
    /product= 4(p63)
    /size= 7441..7440
    /tag= dc
    /product= 4(p63)
    /size= 7441..7440
    /tag= dd
    /product= 4(p63)
    /size= 7441..7440
    /tag= de
    /product= 4(p63)
    /size= 7441..7440
    /tag= df
    /product= 4(p63)
    /size= 7441..7440
    /tag= dg
    /product= 4(p63)
    /size= 7441..7440
    /tag= dh
    /product= 4(p63)
    /size= 7441..7440
    /tag= di
    /product= 4(p63)
    /size= 7441..7440
    /tag= dj
    /product= 4(p63)
    /size= 7441..7440
    /tag= dk
    /product= 4(p63)
    /size= 7441..7440
    /tag= dl
    /product= 4(p63)
    /size= 7441..7440
    /tag= dm
    /product= 4(p63)
    /size= 7441..7440
    /tag= dn
    /product= 4(p63)
    /size= 7441..7440
    /tag= do
    /product= 4(p63)
    /size= 7441..7440
    /tag= dp
    /product= 4(p63)
    /size= 7441..7440
    /tag= dq
    /product= 4(p63)
    /size= 7441..7440
    /tag= dr
    /product= 4(p63)
    /size= 7441..7440
    /tag= ds
    /product= 4(p63)
    /size= 7441..7440
    /tag= dt
    /product= 4(p63)
    /size=
```







FN W0971123-11. 7766 bases\*  
PD 28-AUG-1997.  
PF 20-FEB-1997: US-02952.  
PR 21-FEB-1996: US-01208.  
PS (EXA ) UNIV TEXAS SYSTEM.  
PI Zhang J, Hardham JM, Howell JK, Norris SJ, Weinscock GM,  
DR WPI: 97-4312/240.  
PT Nucleic acid encoding variable major protein-like peptide of  
PT Borrelia or useful for recombinant production of Wp-like protein or  
PT Borrelia or useful for recombinant production of Wp-like protein or  
PT Claim 26: Page 101-105; 110pp; English.  
CC This DNA sequence comprises the variable major protein (VMP)-like  
CC sequence (vls) locus of Borrelia burgdorferi. An infectivity  
CC related 78-bp linear plasmid, pB821a, of B. burgdorferi B1 was  
CC isolated from a 15 silent vls cassette (vls2-vls6), and an  
CC which consists of 15 silent vls cassettes (vls2-vls6), and an  
CC expressed vlsE gene (see T85043). Portions of several of the 15  
CC silent vls cassette sequences, located approx. 500 bp upstream of  
CC infect recombinant to the vlsE gene, are located in the vlsE gene  
CC lipoprotein (see M22676) and hence immune evasion. Long-term  
CC survival and pathogenesis in the mammalian host. Nucleic acids,  
CC including vls2-vls6 sequences, encoding Wp-like proteins, can be  
CC used for recombinant production of Wp-like proteins, or for  
CC the diagnosis of Lyme disease. (claims), and may also have  
CC application in DNA vaccines. (claims), and may also have  
SQ Sequence 7766 BP; 2029 A; 867 C; 3011 G; 1958 T;

Query Match 1.33; Score 14; DB 35; Length 7766;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3400 gagagactgagga 3413  
CP 1013 GAGAGACTGAGGA 1000

RESULT 563  
AC Q84650: standard; DNA: 7791 BP.  
AC Q84650: standard; DNA: 7791 BP.  
DT 01-DEC-1995 (first entry)  
DE Human neuronal calcium channel subunit alpha 1A-2.  
DE Calcium channel subunit; antagonist; agonist; diagnosis;  
KW Leber's Reticular Syndrome; as.  
OS Homo sapiens. Location/Qualifiers  
FH Key 237..7037  
FT cds /\*tag= a  
FT W09504822-A.

PD 16-FEB-1995. U09230.  
PR 11-AUG-1994: US-105516.  
PR 05-NOV-1993: US-149097.  
PA (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.  
PI Ellis SP Gillespie A, Harpold WM, Mocue AP, Williams ME;  
DR P-PSDB: R71007.  
PT DNA encoding human calcium channel sub-unit(s) - used for  
PT developing prods. for studying calcium channels, e.g. for  
PT obtaining agonists and antagonists  
PT The primary transcript of the alpha 1A subunit gene is  
CC alternatively spliced to yield at least two variant mRNAs.  
CC One form, alpha 1A-1 is given in Q84659/R71007, and the other,  
CC alpha 1A-2 is given in Q84660/R71008. Alpha 1A-2 differs from  
CC alpha 1A-1 in that it lacks a 5' nt sequence. This deletion shifts the reading frame and  
CC introduces a translation termination codon resulting in an  
CC alpha 1A-2 coding sequence that encodes a shorter alpha 1A  
CC subunit than that encoded by alpha 1A-1. DNA encoding alpha  
CC subunit that has been deposited in the ATCC under accession  
CC no. 75293. The DNA is such a phase includes the DNA fragment  
CC containing the translation termination codon resulting in an  
CC conditions of high stringency to DNA encoding alpha 1A DNA but not  
CC to DNA encoding alpha 1B.  
SQ Sequence 7791 BP; 1675 A; 2436 C; 2258 G; 1422 T;

Query Match 1.33; Score 14; DB 15; Length 7791;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5363 catctatgcatca 5376  
QY 342 CATCTATGCCATCA 355

RESULT 564  
AC Q84659: standard; DNA: 7808 BP.  
AC Q84659: standard; DNA: 7808 BP.  
DT 01-DEC-1995 (first entry)  
DE Human neuronal calcium channel subunit alpha 1A-1.  
DE Calcium channel subunit; antagonist; agonist; diagnosis;  
KW Leber's Reticular Syndrome; as.  
OS Homo sapiens. Location/Qualifiers  
FH Key 237..7769  
FT cds /\*tag= a  
FT W0971123-11.

FT misc\_difference 7035..7039  
FT /\*tag= b  
FT /\*note= 'not present in alpha 1A-2'  
PD 16-FEB-1995.  
PF 11-AUG-1994: U09230.  
PR 11-AUG-1993: US-105536.  
PR 05-NOV-1993: US-149097.  
PS (EXA ) UNIV TEXAS SYSTEM.  
PI Ellis SP Gillespie A, Harpold WM, Mocue AP, Williams ME;  
DR WPI: 95-090900/12.  
DR P-PSDB: R71007.  
PT DNA encoding human calcium channel sub-unit(s) - used for  
PT developing prods. for studying calcium channels, e.g. for  
PT obtaining agonists and antagonists  
PT The primary transcript of the alpha 1A subunit gene is  
CC alternatively spliced to yield at least two variant mRNAs.  
CC One form, alpha 1A-1 is given in Q84659/R71007, and the other,  
CC alpha 1A-2 is given in Q84660/R71008. Alpha 1A-2 differs from  
CC alpha 1A-1 in that it lacks a 5' nt sequence. This deletion shifts the reading frame and  
CC introduces a translation termination codon resulting in an  
CC alpha 1A-2 coding sequence that encodes a shorter alpha 1A  
CC subunit than that encoded by alpha 1A-1. DNA encoding alpha  
CC subunits can be isolated using all or a portion of the DNA  
CC having sequence Q84661. Q85659 or Q84660 or DNA obt'd. from  
CC the phage lysate of an E. coli host contg. DNA encoding an alpha  
CC subunit that has been deposited in the ATCC under accession  
CC no. 75293. The DNA is such a phase includes the DNA fragment  
CC containing the translation termination codon resulting in an  
CC conditions of high stringency to DNA encoding alpha 1A DNA but not  
CC to DNA encoding alpha 1B.  
SQ Sequence 7808 BP; 1680 A; 2441 C; 2265 G; 1422 T;

Query Match 1.33; Score 14; DB 15; Length 7808;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5363 catctatgcatca 5376  
QY 342 CATCTATGCCATCA 355

RESULT 565  
ID W74884: standard; cDNA: 7892 BP.  
AC W74884: 1998 (first entry)  
DE Miniature swine retrovirus cDNA.  
DE Retrovirus; porcine; GAG protein; POL protein; ENV protein; donor;  
KW xenotransplantation; infectious; provirus; organ transplant; donor;

KW activated virus; PCR; as.  
FH Key 585..2159  
FT cds Location/Qualifiers  
FT mat\_peptide /\*tag= b  
FT /\*note= 'putative GAG protein'  
FT CDS /\*tag= d  
FT mat\_peptide 2307..5741  
FT /\*tag= d  
FT CDS /\*note= 'putative POL protein'  
FT /\*tag= f  
FT mat\_peptide 5620..7533  
FT /\*note= 'putative ENV protein'  
W0971123-11.  
PD 19-JUN-1997.  
PF 13-DEC-1996; U19680.  
PR 14-DEC-1995; US-572645.  
PS (GHEC ) GEN HOSPITAL CORP.  
PI WPI: 97-322804/30.  
DR P-PSDB: W32096-W32098.  
PT New nucleic acid from porcine retro-viruses - used for detecting  
PT viruses in transplant or other tissue and for assessing risk of  
PT infection of recipient  
PT Claim 22: Fig 3; 12pp; English.  
CC This cDNA sequence represents a porcine retrovirus from miniature swine  
CC containing the putative coding regions for viral GAG, POL and ENV  
CC proteins. The sequence and PCR fragments generated from it  
CC were used to detect the presence of  
CC porcine retroviruses prior to xenotransplantation. Transplantation can  
CC increase the likelihood of retroviral activation if intact and infectious  
CC proviruses are present. The porcine retroviral sequence can be used to  
CC generate probes to determine the level (e.g. copy number) of intact  
CC proviruses in transplant tissue. The presence of intact proviruses in  
CC xenograft transplantation donors can be used to detect mutations,  
CC genetic lesions or viral recombinants and also to determine the  
CC histological localization of activated retrovirus. Using Polymerase Chain  
CC Reaction DNA Quantitation (PCR) on blood mononuclear cells, infectivity  
CC of retroviruses can be determined. The presence of intact proviruses in  
CC donors without intact porcine retroviral sequences or a lower copy number  
CC of viral elements could be selected.  
SQ Sequence 7892 BP; 2191 A; 1915 C; 1980 G; 1806 T;

Query Match 1.33; Score 14; DB 15; Length 7892;  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;







Page 570

Page 572



[illegible]

```

Ft Ft      /*tag= 7
Ft Ft      /note="encodes zinc-finger"
Ft Ft      misc_rna
Ft Ft      2184..2262
Ft Ft      /*tag= 6
Ft Ft      /*tag= t
Ft Ft      /note="encodes zinc-finger"
Ft Ft      misc_rna
Ft Ft      3325..3420
Ft Ft      /*tag= 8
Ft Ft      /*tag= g
Ft Ft      /note="encodes zinc-finger"
Ft Ft      misc_rna
Ft Ft      4350..4455
Ft Ft      /*tag= t
Ft Ft      /number= 10
Ft Ft      /note="encodes zinc-finger"
Ft Ft      misc_rna
Ft Ft      4956..5058
Ft Ft      /*tag= u
Ft Ft      /number= 11
Ft Ft      /note="encodes zinc-finger"
Ft Ft      misc_rna
Ft Ft      5509..5604
Ft Ft      /*tag= y
Ft Ft      /number= 12
Ft Ft      /note="encodes zinc-finger"
Ft Ft      misc_rna
Ft Ft      6100..6150
Ft Ft      /*tag= 4
Ft Ft      /number= 13
Ft Ft      /note="encodes zinc-finger"
Ft Ft      misc_rna
Ft Ft      7152..7182
Ft Ft      /*tag= 3
Ft Ft      /number= 14
Ft Ft      /note="encodes zinc-finger"
Ft Ft      misc_rna
Ft Ft      8042..8043
Ft Ft      /*tag= 15
Ft Ft      /number= 15
Ft Ft      /note="encodes zinc-finger"
Ft Ft      BP-4497229.A
Ft Ft      37-MAY-1991; 310334.
Ft Ft      PR 07-NOV-1991; 310334.
Ft Ft      PR 07-NOV-1990; JP-304412.
Ft Ft      PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
Ft Ft      WP 93-176828/22.
Ft Ft      DR P-7503; R23963.
Ft Ft      DR P-7503; R23963.
Ft Ft      PR DNA encoding protein binding to alpha-fetoprotein gene enhancer -
Ft Ft      useful for prom. of biological active protein
Ft Ft      CC This DNA sequence is as the sequence given in Q24928 except that a
Ft Ft      CC cytosine at nucleotide 7508 is replaced by thymine (claim 4). The
Ft Ft      CC cytosine at nucleotide 7509 may also be replaced by other
Ft Ft      CC protein is replaced by other amino acids
Ft Ft      CC

```

This sequence codes for the protein that specifically binds to the enhancer of the alpha-fetoprotein gene. This DNA may be useful for the prodn. of biologically active proteins by inserting this DNA into expression vector in which a gene for the biologically active protein is placed under the control of alpha-fetoprotein gene enhancer and promoter.

8601 BP: 2301 A: 2628 C: 2114 G: 1558 T:  
SQ Squares

Query Match: 1.38; Score 14; DB #: Length 8601;  
Percent Similarity 100.0%; Pred. Res: 1745-02; o, Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0;

DB 5943 cygcgataacagc 5956  
DY 1067 cgcgaatcacgc 1075

```

Pt /note="edminine missing in lambda 489"
Pt 3820..3999
Pt /tag= g
Pt /number= 1
Pt /note="encodes homeodomain"
Pt 4000..4249
Pt /tag= h
Pt /number= 2
Pt /note="encodes homeodomain"
Pt 4250..4367
Pt /tag= i
Pt /number= 3
Pt /note="encodes homeodomain"
Pt 4368..4402
Pt /tag= j
Pt /number= 4
Pt /note="encodes homeodomain"
Pt 4403..4426
Pt /tag= k
Pt /number= 5
Pt /note="encodes zinc-finger"
Pt 4427..4494
Pt /tag= l
Pt /number= 2
Pt /note="encodes zinc-finger"
Pt 4495..4738
Pt /tag= m
Pt /number= 3
Pt /note="encodes zinc-finger"
Pt 4739..4827
Pt /tag= n
Pt /number= 6
Pt /note="encodes zinc-finger"
Pt 4828..4857
Pt /tag= o
Pt /number= 5
Pt /note="encodes zinc-finger"
Pt 4858..4878
Pt /tag= p
Pt /number= 6
Pt /note="encodes zinc-finger"
Pt 4879..4909
Pt /tag= q
Pt /number= 7
Pt /note="encodes zinc-finger"
Pt 4910..4922
Pt /tag= r
Pt /number= 8
Pt /note="encodes zinc-finger"
Pt 4923..4940
Pt /tag= s
Pt /number= 8

```



[illegible]

PT	misc_signal	/label= CRE	
PT		1055..1066	
PT		/tag= k	
PT		/label= BP-IT-beta	
PT	tata_signal	1078..1078	beta-specific cis-acting element*
PT		/tag= l	
PT	exon	1102..1170	
PT		/tag= s	
PT		/number= I	
PT	repeat_unit	1405..1629	
PT		/tag= n	
PT		/rpt_type= DIRECT	
PT		/note= porcine-specific repetitive element PRF-1*	
PT	exon	1909..1970	
PT		/tag= o	
PT		/number= II	
PT		/domain_start= 1924..1926	
PT	exon	2495..2550	
PT		/tag= p	
PT		/number= III	
PT	exon	3701..3902	
PT		/tag= r	
PT		/number= IV	
PT	repeat_unit	4022..4044	
PT		/tag= z	
PT		/note= PRF-1-like element*	
PT	repeat_unit	4465..4716	
PT		/tag= s	
PT		/rpt_type= DIRECT	
PT		/domain_start= 4701..4704	
PT	exon	5001..5241	porcine-specific repetitive element PRF-1*
PT		/tag= t	
PT		/number= v	
PT	exon	6474..6604	
PT		/tag= u	
PT		/number= VI	
PT	exon	7384..7596	
PT		/tag= v	
PT		/number= VII	
PT	misc_signal	7935..7952	
PT		/tag= y	
PT		/label= UA-motif	
PT	polya_signal	8164..8170	
PT		/tag= w	
PT	repeat_unit	8614..8760	
PT		/tag= y	
PT		/rpt_type= DIRECT	
PT		/note= porcine-specific repetitive element PRF-1*	
PT	NC9301929-A		
PT	17-OCT-1994		
PD			



```

4987. 5210
4988. /product= green fluorescent protein
4989. 5220. 5449
4990. /stage g
4991. /product= Shiva-1
4992.
4993. W05640879-Al.
4994. 19-DIC-1995.
4995. 07-JUN-1995; U10041-.
4996. 07-JUN-1995; U5-480454.
4997. 07-JUN-1995; U10041-Strk.
4998. Morrey J, Reed NA, White KL;
4999. WP1; 97-087059/08.
5000.
5001. Expression cassettes encoding amphipathic peptide(s) - used for
5002. e tumor; disease-resistant animals or for the treatment of diseases
5003.
5004. Clin 11; Page 43-48; 75pp. English.
5005. A DNA cassette (W05055) is used to express the amphipathic (lytic)
5006. peptide Shiva-1 in the milk of a transgenic animal. The Shiva-1
5007. peptide is a 100 amino acid peptide that encodes a green fluorescent
5008. regulatory elements. The cassette also encodes a green fluorescent
5009. protein that acts to inhibit the toxic activity of the amphipathic
5010. peptide and to facilitate detection of gene expression and protein
5011. expression. The cassette also encodes a green fluorescent protein
5012. beta casein 3' regulatory elements. Other intramammary-regulated
5013. cassettes can be used to produce disease-resistant animals or to
5014. treat diseases a. e. tumors.
5015. Sequence 8797 BP; 2674 A; 1814 C; 1806 G; 2503 T;
5016.
5017. Query Match 1.3%; Score 14; DB 37; Length 8797;
5018. Best Local Similarity 100.0%; Pred.No. 3.47e+02;
5019. Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
5020.
5021. Db 3689 cctcttggtctctt 3702
5022. |||||
5023. QY 632 TCTTGAGTCTCTT 645
5024.
5025. RESULT 577
5026. ID T03104 standard; DNA: 9108 BP.
5027. AC T03104.1995 (first entry)
5028. DE Plasmid PRK5.tkl-1.1 encoding sal s-1.
5029. KW Protein tyrosine-kinase; PTK; SAL-S1; agonist; cell growth;
5030. differentiation; PRK5.tkl-1.1. es.
5031. OS Chimeric synthetic.
5032. GN Chimeric synthetic.
5033. PN W09570681-Al.
5034. PP 12-OCT-1995.
5035. FD 04-APR-1995; D04228.
5036.

```

CC	(see also O64176), from a heterologous virus so the NANBIV protein
CC	may be expressed by transformed insect cells. The expressed protein
CC	is used as antigen for the development of anti-NANBIV antibodies.
CC	Cc These antigens may be used in a hepatitis C vaccine.
SQ	Sequence 9391 BP; 1862 A; 2836 C; 2699 G; 1995 T;
	Query Match 13%; Score 14; DP 11; Length 9391,
	Best Local Similarity 100.0%; Pred. No. 3/47+02;
	Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dd	8533 tctgtctatctggg 8546 
Cp	1075 ttctgcattatctggc 1062
RESULT 579	ID Q38959 standard; cDNA; 9391 BP.
AT	QC Q38959;
DC	29-JUL-1993 v(first entry)
KN	HCV; detection: antigen; vaccine/recombinant, aa.
OS	Hepatitis C virus.
FH	Key Location/Qualifiers
FE	cds 305..3349
CD	/cds = a
JD	JO5068563-A.
Pd	23-MAR-1993.
PN	FF 17-JUL-1991; 203984
PF	17-JUL-1991; 203984
PA	(MAG) KAKUGU OTSUI KESSEI RYOKO.
WP:	93-130639/16.
DR	P-PDSB, R35207.
DT	Genetic engineering.
PT	useful for detecting HCV infection, esp. in Japan, and as vaccine
PT	against HCV
PS	Clinical; Page 6-17; 17pp; Japanese.
CS	NHAN was extracted from the plasma of Japanese patients whose HBs
CC	Ag was detected positive by enzyme-linked immunosorbent assay
CC	from the RNA and cloned into lambda gt10 and screened using HCV
CC	infected chimpanzee plasma to isolate HCV clones. The DNA sequence
CC	is useful in detection of HCV virus. The polypeptide it produces
CC	is used in detecting antibody against the protein of HCV vaccine.
SQ	Sequence 9391 BP; 1861 A; 2840 C; 2698 G; 1994 T;
	Query Match 13.3%; Score 14; DB 6; Length 9391;
	Best Local Similarity 100.0%; Pred. No. 3/47+02;
	Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dd	8533 tctgtctatctggg 8546 
Cp	1075 ttctgcattatctggc 1062



[illegible]

```

US-08-887-977-r.mg

Tue Nov 17 08:55:36 1998

PA 15-FEB-1995: US-388909.
PA (FRUSH ) US DEPT HEALTH & HUMAN SERVICES.
PI (FRUSH ) G. Gallo RC, Gartner S, Lori FC, Markham PD;
PI WPI: 97-011206/01 MS;
DR WPI: 97-011206/01 MS;
DR P-FSMB: M11579.
DR P-FSMB: M11579.
PT New isolated envelope protein of HIV-1 strain BA-L and recombinant
PT equivalents useful as immunogens for vaccines and antibody prodn.,
PT especially for HIV-1 strains that are resistant to AZT.
PS Envelope protein: amino acid residues 1-232.
PS Envelope: 218-232.
PS Envelope: 218-232.
CC The permuted circular unintegrated viral DNA representing the
CC complete HIV-1(10N) genome was cloned into the EcoRI site of lambda
CC vector lambda B DNA from total DNA of B9 cells producing HIV-1(10N)
CC genome. Lambda B DNA from total DNA of B9 cells producing HIV-1(10N)
CC genome was also cloned into the EcoRI site of lambda B DNA. The
CC insert having the present sequence 1773 C: 2344 G: 2160 T:
SQ Sequence 9739 BP: 3462 A: 3462 G: 2160 T:

Query Match 1.3% Score 14: DB 27; Length 9739;
Seq. Id. 100%; Seq. Id. 100%; Seq. Id. 100%;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 8908 gsgagcagctcccc 8921
CP 379 GAGCAGCAGCTCCG 366

RESULT 582
ID Q14751 standard; DNA; 9739 BP.
ID Q14751 standard; DNA; 9739 BP.
DE 05-FEB-1992 (first entry)
DE HIV-1(10N) env protein-encoding sequence.
KW human immunodeficiency virus; United States; MN isolate; AIDS;
KW envelope protein; aa, virus-1 (10N).
PR Genbank
PR Location/Qualifiers
FW 6240..8810
cde
FT /tag= a
FT /product= env
FT U57599491..N
PD 15-OCT-1991.
PD 15-OCT-1991.
PD 17-OCT-1990: 188330.
PR 17-OCT-1990: US-599493.
PA (USHS ) NAT INST OF HEALTH.
PA (USHS ) NAT INST OF HEALTH.
DR WPI: 91-146752/47.
DR P-FSMB: R14903.
DR P-FSMB: R14903.
PT US HIV-1 isolates MN-STI and BA-L, ENV protein and DNA - are
PT useful in therapeutics, vaccines and diagnostic tests
PT complete HIV-1(10N) genome was cloned into the EcoRI site of lambda
PT vector lambda B DNA from total DNA of B9 cells producing HIV-1(10N)
PT genome. Lambda B DNA from total DNA of B9 cells producing HIV-1(10N)
PT genome was also cloned into the EcoRI site of lambda B DNA. The
PT insert having the present sequence 1773 C: 2344 G: 2160 T:
SQ Sequence 9739 BP: 3462 A: 3462 G: 2160 T:

```

```

(INSP) Inst Pasteur(AlZV).D;
PA Alizon M, Montagnier L, Gueudet D, Clavié F, Sonigo P, Guyader M,
PI Tellois P, Chakrabarti L, Desrosiers R;
P2 P-867890070, p187783, p817694, p81758, p81757, p81756,
DR p-86759, p81760, p81761, p81762;
DD New peptides(s) with immunological properties of HIV-2 envelope protein -
DT having the structure of simian immune deficiency virus proteins,
PT claim 23; fig 13-86pp; French;
PP n08060 is contained in lambda giv-1 which was deposited on 15/4/87 at
CC the CNRS under number I-659 (gsv-1.1.1 and I-659 (gsv-1.1.2). It is the
CC cDNA to sv1 genomic RNA. Recombinant DNA contg. all or part of the cDNAs
CC of HIV-1 and HIV-2 are used as hybridisation probes. New peptides which have
CC immunological properties in common with those of the peptide skeleton of
CC the envelope protein of HIV-2 and also have a peptide structure in common
CC with the envelope protein of HIV-1 are conjugated with a fluorescent antibody
CC conjugates new. The peptides and a kit to detect HIV-1 immunogenic
CC fluids are sent. The peptides are useful for in vitro diagnosis of HIV-2
CC infection and some of them can be used as components of immunogens and
CC vaccines against HIV. Antibodies raised against them can be used for
CC treatment of HIV infected patients.
SQ Sequence 9601 BP; 3277 A; 1809 C; 2178 G; 2136 T; 1 Others:
Query Match Similarity 10.3%; Score 14; DB 1; Length 9601;
Local 14; Conservative 0; Gap=0; Re-Op; Indels 0; Gaps 0;
Matches 14; Mismatches 0;
Db 8537 agctccgcggcgtt 8550
QY 997 AUGCTGCTGGCGT 1010
RESULT 581
AC ID 1585858 standard. cDNA. 9739 BP.
AD T38549;
AK T38549;
AS 25-MAR-1987 (first entry)
DE Human immunodeficiency virus type 1 strain MN-PH1 genome.
DI Acquired immune deficiency syndrome; AIDS; envelope protein.
DS Human immunodeficiency virus type 1 strain MN-PH1.
OS Key Location/Qualifiers
FE Key 6240..8810
FT /product= envelope_protein
FT /
FT /
FT U85576000-A.
PD 19-NOV-1986. 589693
PD 17-OCT-1990. US-594491.
PR 25-FEB-1993; US-022835.
```

[illegible]







Tue Nov 17 08:55:26 1998

KW		gene therapy; tumour; da.
OS		Chimeric Bos taurus;
OS		Chimeric Mus sp.;
OS		Chimeric Synthetic.
FT	key	Accession/Qualifiers
FT	S'attr	127-1800
FT		/tag= a
FT	tata_signal	1766..1773
FT		/tag= b
FT	exon	1801..1823
FT		/tag= c
FT		/note= beta casein exon 1'
FT	exon	3780..3832
FT		/tag= d
FT		/note= beta casein exon 2'
FT	exon	4587..4590
FT		/tag= e
FT		/note= portion of beta casein exon 3'
FT	mat_peptide	4587..5310
FT		/tag= f
FT		/note= green fluorescent protein
FT	mat_peptide	5320..5449
FT		/tag= g
FT		/product= Shiva-1
FT	polya_signal	7630..7635
FT		/standard_name= bovine beta casein 3' region, in exon g
FN		W09640879-A1.
PN		D9-DIC-1986.
PR		03-JUN-1995: US 410041.
PR		07-JUN-1995: US 480454.
PA		(UTAH) UNIV UTAH STATE,
PI		Morrey J., Reed WJ, KL;
PI		Fri.: 97-08/05/98
DR		Protein encoding amphibian peptide(s) - used for treatment of disease-resistant animal or for the treatment of diseases

[illegible]

Tue Nov 17 08:55:26 1998

85-08-887-977-9. FBI

**Expression cassettes encoding amphiphilic peptide(s) - used for prodn. of disease-resistant animals or for the treatment of diseases e.g. tumours**

Cc 1: Page 3643; 75pp; English.

Cc 2: Claim 11; Page 0095; 5pp; English.

Cc 3: The Shiva-1 gene expresses the amphiphilic (lytic) peptide, Shiva-1, in the milk of a transgenic animal. The Shiva-1 gene is placed under control of bovine milk-specific beta casein 5' and 3' regulatory elements. The cassette also encodes a green fluorescent protein (GFP) gene. The GFP gene is placed under the amphiphilic peptide and to facilitate detection of gene expression and protein purification. Another cassette (see also US0856) lacks beta casein 3' regulatory elements. Other, interleukin-regulated cassettes can be used to produce disease-resistant animals or to produce disease-resistant cells.

Cc 4: Claim 1; Page 0093; 5pp; English.

Cc 5: Sequence 11093 BP; 13392 A; 2234 C; 2152 G; 3315 T; 3315 C.

```

Query Match      1.3%  Score 15:  DB 27:  Length 11093;
Best Local Similarity 100.0%  Pred. No. 0.37e+01;
Matches 15;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Db 7537 aaatcatcatgaacat7551
    |||||
Cp 669 aaatcatcatgaacat7551

```

Accession	Key	Location/Qualifiers
BSR001	539	
IC	T8481	standard; cDNA to mRNA, 11293 BP.
AC	T8481.1	
DE	21-FEB-1998	(first entry)
DE	Human breast and ovarian cancer susceptibility gene BRCA2.	
DE	BRCA2 gene, human breast and ovarian cancer suppressor gene	
OS	Human sapiens.	
FS	Key	
FT	CDS	Location/Qualifiers
FT		137..10453
FT		/tag= a
PN	W09730108-A1.	
PD	21-AGC-1997.	
PF	19-FEB-1997;	003340.53
PA	(UNIV) UNIV WASHINGTON.	
PA	(UNIV) UNIV WASHINGTON.	
PI	Clair-Rking M, Holt JT, Jensen RA, Jettom TL, Page DL;	
PI	Robinson, Benson CL, Szabo CI, Thompson ME;	
PI	P-FSDB: W43287.46.	
DR	P-FSDB: W43287.46.	
DR	BRCA1 and BRCA2 tumour suppressor gene products - useful to inhibit	
PT	breast and ovarian cancer cell growth and tumorigenesis, or treat	
PT	genetic therapy.	
CC	Genbank entry 7216746.	
CC	This sequence comprises tumour suppressor gene BRCA2 cDNA. BRCA2	

Tue Nov 17 08:55:26 1998

```

Query Match      1.38;      Score 14;  DB 27;  Length 11083;
Best Local Similarity 100.0%;  Pred. No. 3,476+02;
Mismatch 14;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

RESULT      588
D          T50855 standard; DNA; 11093 bp.
P          T50855:
I          30-MAR-1987 (first entry)
O          DNA cassette for amphipathic peptide production in milk.
C          Amphipathic peptide:milk; transgenic animal; beta-casein; promoter;
E          gene therapy; tumour; da.
S          Chimeric Bos taurus;
T          ChimERIC Rna sp.;
V          ChimERIC synthetic.
W          Location/Qualifiers
X          137..1800
Y          5'utr
Z          tata_signal
A          /tag= a
B          1765..1773
C          /tag= b
D          exon
E          1801..1823
F          /note= "beta casein exon 1"
G          exon
H          3780..3832
I          /tag= d
J          /note= beta casein exon 2"
K          exon
L          4530..4590
M          /tag= e
N          /note= "portion of beta casein exon 3"
O          mat_peptide
P          4587..5310
Q          /tag= f
R          product= green fluorescent protein
S          5320..5469
T          /tag= g
U          poly_a_signal
V          /product= Shiva-1
W          7830..7695
X          /tag= h
Y          exon
Z          8
A          WO9640879-A1.
B          D
C          19-DEC-1986.
D          O
E          07-JUN-1986; 010041
F          CHIMERIC RNA SP.
G          (UPAR ) ENTRY
H          Morrey J, Reed Wv, White KL;
I          MPI; 97-087059/08.

```

3 WPI; 97-087059/08.

THE NOV 17 08:55:26 1998

118-08-887-977-9 MC

encodes a protein (see W31287), that acts as a growth inhibitor of breast and ovarian cancers. DNA encoding the BRCA1 protein can thus be used in gene therapy methods for the treatment of breast and ovarian cancers. Purified BRCA1 protein can also be used in gene therapy methods for the treatment of breast and ovarian cancers. The BRCA1 gene can also be used to identify the BRCA2 receptor, and hence to identify ligands that can be used as protein-mimetics which act on the receptor for use in breast and ovarian cancer therapy. The BRCA1 gene (see T84840) and BRCA1 protein (see W31288) have also been characterized. Methods are provided for identifying and isolating BRCA1 gene sequences for treating or preventing breast and (gene-linked hereditary or sporadic) ovarian cancers using BRCA1 and BRCA2 proteins and genes.

Sequence 11763 BP: 4068 A, 1971 C, 2091 G, 3135 T;

```

Query Match      1.38;      Score 14; DB 35; Length 11283;
Best Local Similarity 100.0%; Pred No 3,47+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
9 7888 tgaagattatttg 7901
| | | | | | | | | | | | | | | |
17 tgaagattatttg 30

```

```

USULT 590
0 759707 standard; cDNA; 11385 BP.
1 10-MAR-1998 (first entry)
2 Human breast cancer susceptibility gene BRCA2;
3 Human breast cancer; susceptibility; gene BRCA2; diagnosis;
4 Screening; treatment; gene therapy; ss.
5 Recd aspin. Location/Qualifiers
6 Key CDS
7 229..10485
8 /tag= a
9 W09723690..1.
10 26-JUN-1997.
11 17-DEC-1996; 019598.
12 28-APR-1996; OS-639501.
13 28-DEC-1995; OS-575579.
14 28-DEC-1995; OS-575579.
15 21-DEC-1995; OS-576559.
16 11-JAN-1996; OS-565351.
17 (HSCR) HSC RES & DEV L.
18 (HSCR) HSC RES & DEV L.
19 (HSCR) HSC RES & DEV L.
20 (TYPE) UNK DESTINATION.
21 (TYPE) UNK DESTINATION.
22 (TYPE) UNK DESTINATION.
23 Couch F, Kamb A, Rommens JM, Simard J, Tavtigian SV,
24 Weber BL,
25 Weber BL,
26 P-RUB: M9211.

```















PT polyA\_signal /\*tag= af  
 FT 10814..10817  
 PT polyA\_signal /\*tag= h  
 FT 10839..10842  
 PT polyA\_signal /\*tag= ah  
 FT 10865..10868  
 PT polyA\_signal /\*tag= ai  
 FT 10891..10894  
 PN J06098656-A  
 PD 12-APR-1994  
 PF 30-MAR-1992: JP-102500  
 PR 10-MAR-1992: JP-102500  
 DR 10-MAR-1992: JP-102500  
 WR 94-15585/15  
 PT Transgenic rice containing the rice starch branch family enzyme -  
 FT used to increase the amylopectin content of albumen  
 PT Claim 1: Page 16-21: 24pp: Japanese  
 CC a rice increase the activity of this enzyme in the plant, thereby  
 CC increasing the content of amylopectin in albumen starch and thus  
 CC enabling efficient mass production of various proteins.  
 CC Sequence 12151 BP: 3269 A; 2470 C; 2518 G; 3891 T;  
 SQ

Query Match  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 7828 agacacacagagga 7841  
 Cp 264 AGACACACAGGGA 251

RESULT 595  
 ID T47715 standard: DNA: 15144 BP.  
 AC T47715:  
 DE 30-MAY-1997 (first entry)  
 KW Osteogenic agent: bone morphogenetic protein-2; BMP-2;  
 KW growth factor; osteoblast; promoter; osteoporosis; fracture repair;  
 KW osteoblastic metastasis; osteosclerosis; therapy; as.  
 PR 05-DEC-1996  
 PF 31-MAY-1996: U08197.  
 PR 02-JUN-1995: US-458434.  
 PR 08-SEP-90: US-458434.  
 PR 08-SEP-90: US-458434.  
 DR 97-034396/03  
 WR 97-034396/03  
 PT System for identifying osteogenic agents that induce prodn. of bone  
 PT morphogenetic protein - is cell contg. reporter gene under control of  
 PT promoter. Also new promoters of BMP-2 and -4 and related  
 PT vectors and cells

PS Example 2: Page 33-42: 76pp: English.  
 CC A DNA sequence (T47715) comprises the mouse morphogenetic protein-2  
 CC gene 5' leader and promoter regions. The gene was isolated from a mouse  
 CC genomic library (86/038) using mouse BMP-2 cDNA as a probe. The  
 CC of the 5' flanking sequence (see also T47713) of the gene revealed  
 CC putative response elements and transcription factor recognition  
 CC sequences. The promoter region of the BMP-2 or BMP-4 (see also  
 CC T47713) gene was used to construct a recombinant expression vector.  
 CC used to identify novel osteogenic agents that stimulate bone cells  
 CC to produce BMP family bone growth factors. Such osteogenic agents  
 CC can be used to promote bone growth and healing.  
 CC Sequence 15144 BP: 4002 A; 3380 C; 3503 G; 4199 T;  
 SQ

Query Match  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 13145 tctgaagatctt 13158  
 Cp 953 TCTGAGATCTT 966

RESULT 596  
 ID T41418 standard: DNA: 18319 BP.  
 AC T41418:  
 DE 1997 (first entry)  
 KW HCV; HCV; vaccine; diagnosis; ds.  
 KW Human cytomegalovirus Toledo strain.  
 PR 05-DEC-1996  
 PF 31-MAY-1996: U08197.  
 PR 02-JUN-1995: US-458434.  
 PR 08-SEP-90: US-458434.  
 DR 97-034396/03  
 WR 97-034396/03  
 PT System for identifying osteogenic agents that induce prodn. of bone  
 PT morphogenetic protein - is cell contg. reporter gene under control of  
 PT promoter. Also new promoters of BMP-2 and -4 and related  
 PT vectors and cells

FT cds /product= UL139  
 FT 4484..4828  
 FT /\*tag= h  
 FT 5088..5175  
 FT /\*tag= i  
 FT 6448..7168  
 FT /product= UL141  
 FT 7353..7631  
 FT /\*tag= x  
 FT /product= UL143  
 FT 8867..9169  
 FT /\*tag= m  
 FT 9450..9803  
 FT /\*tag= n  
 FT /product= UL146  
 FT 9888..10347  
 FT /product= UL147  
 FT 10646..11596  
 FT /\*tag= p  
 FT 12712..12805  
 FT /\*tag= q  
 FT /product= UL132  
 FT 13109..13753  
 FT /\*tag= f  
 FT 15756..16124  
 FT /\*tag= s  
 FT /product= UL149  
 FT complement (15874..17802)  
 FT /product= UL150  
 FT complement (17289..18299)  
 FT /\*tag= u  
 FT 7353..7631  
 FT /\*tag= v  
 FT /note= "base 1795 is given as n in the  
 FT specification"  
 FT 8105  
 FT /\*tag= w  
 FT /note= "base 18105 is given as n in the  
 FT specification"

FT misc\_difference 18154  
 FT /\*tag= x  
 FT /note= "base 18154 is given as n in the  
 FT specification"  
 FT 18154  
 FT /\*tag= y  
 FT /note= "base 18160 is given as n in the  
 FT specification"  
 FT 18160  
 FT W05500.97.A1  
 FT 03-DEC-1995  
 PR 26-MAR-1996: U04100.  
 PR 31-MAR-1995: US-414926.  
 PA (AVIR-) AVIRON.  
 DR 96-032555 A1  
 DR W05500.97.A1  
 DR P-PSDB: W05500.1  
 DR P-PSDB: W05502.1  
 DR P-PSDB: W05503.1  
 DR P-PSDB: W05504.1  
 DR P-PSDB: W05505.1  
 DR P-PSDB: W05506.1  
 DR P-PSDB: W05507.1  
 DR P-PSDB: W05508.1  
 DR P-PSDB: W05509.1  
 DR P-PSDB: W05510.1  
 DR P-PSDB: W05511.1  
 DR P-PSDB: W05512.1  
 DR P-PSDB: W05513.1  
 DR P-PSDB: W05514.1  
 DR P-PSDB: W05515.1  
 DR P-PSDB: W05516.1  
 DR P-PSDB: W05517.1  
 DR P-PSDB: W05518.1  
 DR P-PSDB: W05519.1  
 DR P-PSDB: W05520.1  
 PT New isolated human cytomegalovirus nucleic acid - from Towne and  
 PT Toledo strains, used to develop prods. for the diagnosis, prevention  
 PT and treatment of human CMV infections  
 PS Claim: Recombinant human cytomegalovirus (HCV) strains  
 CC A novel nucleic acid (T41418) from the Toledo strain of human  
 CC cytomegalovirus (HCV) includes 21 open reading frames, 19 of which  
 CC encode novel proteins (W05502-20), while 2 encode proteins (W05500-  
 CC W05501). The DNA of fragments of it can be used to direct prodn. of  
 CC recombinant HCV strains. The recombinant HCV strains can be used to  
 CC prodn., and also provide probes useful for the diagnosis of HCV  
 CC infection. Recombinant HCV AD169 or Towne strains carrying Toledo  
 CC sequences can be used in immunising compositions. Novel DNA (see  
 CC T41418) was also obt. from HEP-2 cells. Sequence: 18318 BP: 4110 A; 5192 C; 4040 T;  
 SQ



Query Match  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Length 18318;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14246 atgaggagtaga 14259  
CP 274 ATCCGAGAGTACGA 261

RESULT 597  
ID T16806 standard; DNA: 20710 BP.  
AC T16806 standard; DNA: 20710 BP.  
DE T16806 standard; DNA: 20710 BP.  
KW Kaposi's sarcoma associated herpesvirus clone K5S.  
KW Kaposi's sarcoma associated herpesvirus; KSHV; gamma-2 herpesvirus;  
OS Kaposi's sarcoma; AIDS; lymphoma; AIDS; vaccine; diagnosis; therapy;  
PT Kaposi's sarcoma associated herpesvirus.  
PN W09606159-A1.  
PP 29-FEB-1996.  
PR 11-AUG-1995; U10194.  
PR 18-AUG-1994; US-222365.  
PR 11-APR-1995; US-420235.  
PA (UTCO ) UNIV COLUMBIA NEW YORK.  
PA (GRAM/) GRANT D E.  
PA (VIEI/) VIELE L PS. Grant DE, Vile L;  
DR WPI: 96-151362/15. Kaposi's sarcoma - also associated  
PT Herpesvirus DNA associated with Kaposi's sarcoma - also associated  
PT vectors and proteins, used in detection and vaccination.  
CC Kaposi's sarcoma associated herpesvirus (KSHV) genome. English.  
CC DNA sequences in Kaposi's sarcoma (KS) tissue. An isolated sequence.  
CC K5S108am (T16824), was used to screen a KS lesion genomic library. A  
CC lambda phage, K5S, was identified and the sequence of its insert  
CC with K5S108am. K5S represents novel infectious agent associated  
CC with AIDS-KS. Kaposi's sarcoma associated herpesvirus (KSHV) genome  
CC fragment has 15 complete ORFs and 2 incomplete ORFs (see also T16807  
CC 23). Portions of the sequence may be incorporated into a vector and  
CC expressed in host cells to produce peptides (see also R93601-17)  
CC used to immunize animals and to produce antisera. The antisera  
CC are used to immunize a subject or used to design therapeutic  
CC antisense and triplex molecules or diagnostic probes, or to raise  
CC transgenic animals.  
SQ Sequence 20710 BP; 4628 A; 5526 C; 5666 G; 4890 T;  
Query Match  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7069 gagctctgatgtt 7083

Query Match  
Best Local Similarity 100.0%; Pred. No. 3.47e+02; Length 18318;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14246 atgaggagtaga 14259  
CP 274 ATCCGAGAGTACGA 261

RESULT 598  
ID T30681 standard; DNA: 20710 BP.  
AC T30681 standard; DNA: 20710 BP.  
DE T30681 standard; DNA: 20710 BP.  
KW Kaposi's sarcoma associated herpesvirus clone K5S.  
KW Kaposi's sarcoma associated herpesvirus; KSHV; gamma-2 herpesvirus;  
OS Kaposi's sarcoma; AIDS; lymphoma; AIDS; vaccine; diagnosis; therapy;  
PT Kaposi's sarcoma associated herpesvirus.  
PN W09606159-A1.  
PP 29-FEB-1996.  
PR 11-AUG-1995; U10194.  
PR 18-AUG-1994; US-222365.  
PR 11-APR-1995; US-420235.  
PA (UTCO ) UNIV COLUMBIA NEW YORK.  
PA (GRAM/) GRANT D E.  
PA (VIEI/) VIELE L PS. Grant DE, Vile L;  
DR WPI: 96-151362/15. Kaposi's sarcoma - also associated  
PT Herpesvirus DNA associated with Kaposi's sarcoma - also associated  
PT vectors and proteins, used in detection and vaccination.  
CC Kaposi's sarcoma associated herpesvirus (KSHV) genome. English.  
CC DNA sequences in Kaposi's sarcoma (KS) tissue. An isolated sequence.  
CC K5S108am (T16824), was used to screen a KS lesion genomic library. A  
CC lambda phage, K5S, was identified and the sequence of its insert  
CC with K5S108am. K5S represents novel infectious agent associated  
CC with AIDS-KS. Kaposi's sarcoma associated herpesvirus (KSHV) genome  
CC fragment has 15 complete ORFs and 2 incomplete ORFs (see also T16807  
CC 23). Portions of the sequence may be incorporated into a vector and  
CC expressed in host cells to produce peptides (see also R93601-17)  
CC used to immunize animals and to produce antisera. The antisera  
CC are used to immunize a subject or used to design therapeutic  
CC antisense and triplex molecules or diagnostic probes, or to raise  
CC transgenic animals.  
SQ Sequence 20710 BP; 4628 A; 5526 C; 5666 G; 4890 T;  
Query Match  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7069 gagctctgatgtt 7083

Query Match  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14246 atgaggagtaga 14259  
CP 274 ATCCGAGAGTACGA 261

RESULT 599  
ID T34137 standard; DNA: 37895 BP.  
AC T34137 standard; DNA: 37895 BP.  
DE T34137 standard; DNA: 37895 BP.  
KW Kaposi's sarcoma associated herpesvirus clone K5S.  
KW Kaposi's sarcoma associated herpesvirus; KSHV; gamma-2 herpesvirus;  
OS Kaposi's sarcoma; AIDS; lymphoma; AIDS; vaccine; diagnosis; therapy;  
PT Kaposi's sarcoma associated herpesvirus.  
PN W09606159-A1.  
PP 29-FEB-1996.  
PR 11-AUG-1995; U10194.  
PR 18-AUG-1994; US-222365.  
PR 11-APR-1995; US-420235.  
PA (UTCO ) UNIV COLUMBIA NEW YORK.  
PA (GRAM/) GRANT D E.  
PA (VIEI/) VIELE L PS. Grant DE, Vile L;  
DR WPI: 96-151362/15. Kaposi's sarcoma - also associated  
PT Herpesvirus DNA associated with Kaposi's sarcoma - also associated  
PT vectors and proteins, used in detection and vaccination.  
CC Kaposi's sarcoma associated herpesvirus (KSHV) genome. English.  
CC DNA sequences in Kaposi's sarcoma (KS) tissue. An isolated sequence.  
CC K5S108am (T16824), was used to screen a KS lesion genomic library. A  
CC lambda phage, K5S, was identified and the sequence of its insert  
CC with K5S108am. K5S represents novel infectious agent associated  
CC with AIDS-KS. Kaposi's sarcoma associated herpesvirus (KSHV) genome  
CC fragment has 15 complete ORFs and 2 incomplete ORFs (see also T16807  
CC 23). Portions of the sequence may be incorporated into a vector and  
CC expressed in host cells to produce peptides (see also R93601-17)  
CC used to immunize animals and to produce antisera. The antisera  
CC are used to immunize a subject or used to design therapeutic  
CC antisense and triplex molecules or diagnostic probes, or to raise  
CC transgenic animals.  
SQ Sequence 20710 BP; 4628 A; 5526 C; 5666 G; 4890 T;  
Query Match  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7069 gagctctgatgtt 7083

Query Match  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14246 atgaggagtaga 14259  
CP 274 ATCCGAGAGTACGA 261

RESULT 599  
ID T34137 standard; DNA: 37895 BP.  
AC T34137 standard; DNA: 37895 BP.  
DE T34137 standard; DNA: 37895 BP.  
KW Kaposi's sarcoma associated herpesvirus clone K5S.  
KW Kaposi's sarcoma associated herpesvirus; KSHV; gamma-2 herpesvirus;  
OS Kaposi's sarcoma; AIDS; lymphoma; AIDS; vaccine; diagnosis; therapy;  
PT Kaposi's sarcoma associated herpesvirus.  
PN W09606159-A1.  
PP 29-FEB-1996.  
PR 11-AUG-1995; U10194.  
PR 18-AUG-1994; US-222365.  
PR 11-APR-1995; US-420235.  
PA (UTCO ) UNIV COLUMBIA NEW YORK.  
PA (GRAM/) GRANT D E.  
PA (VIEI/) VIELE L PS. Grant DE, Vile L;  
DR WPI: 96-151362/15. Kaposi's sarcoma - also associated  
PT Herpesvirus DNA associated with Kaposi's sarcoma - also associated  
PT vectors and proteins, used in detection and vaccination.  
CC Kaposi's sarcoma associated herpesvirus (KSHV) genome. English.  
CC DNA sequences in Kaposi's sarcoma (KS) tissue. An isolated sequence.  
CC K5S108am (T16824), was used to screen a KS lesion genomic library. A  
CC lambda phage, K5S, was identified and the sequence of its insert  
CC with K5S108am. K5S represents novel infectious agent associated  
CC with AIDS-KS. Kaposi's sarcoma associated herpesvirus (KSHV) genome  
CC fragment has 15 complete ORFs and 2 incomplete ORFs (see also T16807  
CC 23). Portions of the sequence may be incorporated into a vector and  
CC expressed in host cells to produce peptides (see also R93601-17)  
CC used to immunize animals and to produce antisera. The antisera  
CC are used to immunize a subject or used to design therapeutic  
CC antisense and triplex molecules or diagnostic probes, or to raise  
CC transgenic animals.  
SQ Sequence 20710 BP; 4628 A; 5526 C; 5666 G; 4890 T;  
Query Match  
Best Local Similarity 100.0%; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7069 gagctctgatgtt 7083



[illegible]

```

FT 1 /tag> g Enzyme used in EPA biosynthetic pathway.
FT 2 /product> 134345
FT 3 12750 134345
FT 4 /tag> h
FT 5 /product> Enzyme used in EPA biosynthetic pathway.
FT 6
FT 7
FT 8
FT 9
FT 10
FT 11
FT 12
FT 13
FT 14
FT 15
FT 16
FT 17
FT 18
FT 19
FT 20
FT 21
FT 22
FT 23
FT 24
FT 25
FT 26
FT 27
FT 28
FT 29
FT 30
FT 31
FT 32
FT 33
FT 34
FT 35
FT 36
FT 37
FT 38
FT 39
FT 40
FT 41
FT 42
FT 43
FT 44
FT 45
FT 46
FT 47
FT 48
FT 49
FT 50
FT 51
FT 52
FT 53
FT 54
FT 55
FT 56
FT 57
FT 58
FT 59
FT 60
FT 61
FT 62
FT 63
FT 64
FT 65
FT 66
FT 67
FT 68
FT 69
FT 70
FT 71
FT 72
FT 73
FT 74
FT 75
FT 76
FT 77
FT 78
FT 79
FT 80
FT 81
FT 82
FT 83
FT 84
FT 85
FT 86
FT 87
FT 88
FT 89
FT 90
FT 91
FT 92
FT 93
FT 94
FT 95
FT 96
FT 97
FT 98
FT 99
FT 100
FT 101
FT 102
FT 103
FT 104
FT 105
FT 106
FT 107
FT 108
FT 109
FT 110
FT 111
FT 112
FT 113
FT 114
FT 115
FT 116
FT 117
FT 118
FT 119
FT 120
FT 121
FT 122
FT 123
FT 124
FT 125
FT 126
FT 127
FT 128
FT 129
FT 130
FT 131
FT 132
FT 133
FT 134
FT 135
FT 136
FT 137
FT 138
FT 139
FT 140
FT 141
FT 142
FT 143
FT 144
FT 145
FT 146
FT 147
FT 148
FT 149
FT 150
FT 151
FT 152
FT 153
FT 154
FT 155
FT 156
FT 157
FT 158
FT 159
FT 160
FT 161
FT 162
FT 163
FT 164
FT 165
FT 166
FT 167
FT 168
FT 169
FT 170
FT 171
FT 172
FT 173
FT 174
FT 175
FT 176
FT 177
FT 178
FT 179
FT 180
FT 181
FT 182
FT 183
FT 184
FT 185
FT 186
FT 187
FT 188
FT 189
FT 190
FT 191
FT 192
FT 193
FT 194
FT 195
FT 196
FT 197
FT 198
FT 199
FT 200
FT 201
FT 202
FT 203
FT 204
FT 205
FT 206
FT 207
FT 208
FT 209
FT 210
FT 211
FT 212
FT 213
FT 214
FT 215
FT 216
FT 217
FT 218
FT 219
FT 220
FT 221
FT 222
FT 223
FT 224
FT 225
FT 226
FT 227
FT 228
FT 229
FT 230
FT 231
FT 232
FT 233
FT 234
FT 235
FT 236
FT 237
FT 238
FT 239
FT 240
FT 241
FT 242
FT 243
FT 244
FT 245
FT 246
FT 247
FT 248
FT 249
FT 250
FT 251
FT 252
FT 253
FT 254
FT 255
FT 256
FT 257
FT 258
FT 259
FT 260
FT 261
FT 262
FT 263
FT 264
FT 265
FT 266
FT 267
FT 268
FT 269
FT 270
FT 271
FT 272
FT 273
FT 274
FT 275
FT 276
FT 277
FT 278
FT 279
FT 280
FT 281
FT 282
FT 283
FT 284
FT 285
FT 286
FT 287
FT 288
FT 289
FT 290
FT 291
FT 292
FT 293
FT 294
FT 295
FT 296
FT 297
FT 298
FT 299
FT 300
FT 301
FT 302
FT 303
FT 304
FT 305
FT 306
FT 307
FT 308
FT 309
FT 310
FT 311
FT 312
FT 313
FT 314
FT 315
FT 316
FT 317
FT 318
FT 319
FT 320
FT 321
FT 322
FT 323
FT 324
FT 325
FT 326
FT 327
FT 328
FT 329
FT 330
FT 331
FT 332
FT 333
FT 334
FT 335
FT 336
FT 337
FT 338
FT 339
FT 340
FT 341
FT 342
FT 343
FT 344
FT 345
FT 346
FT 347
FT 348
FT 349
FT 350
FT 351
FT 352
FT 353
FT 354
FT 355
FT 356
FT 357
FT 358
FT 359
FT 360
FT 361
FT 362
FT 363
FT 364
FT 365
FT 366
FT 367
FT 368
FT 369
FT 370
FT 371
FT 372
FT 373
FT 374
FT 375
FT 376
FT 377
FT 378
FT 379
FT 380
FT 381
FT 382
FT 383
FT 384
FT 385
FT 386
FT 387
FT 388
FT 389
FT 390
FT 391
FT 392
FT 393
FT 394
FT 395
FT 396
FT 397
FT 398
FT 399
FT 400
FT 401
FT 402
FT 403
FT 404
FT 405
FT 406
FT 407
FT 408
FT 409
FT 410
FT 411
FT 412
FT 413
FT 414
FT 415
FT 416
FT 417
FT 418
FT 419
FT 420
FT 421
FT 422
FT 423
FT 424
FT 425
FT 426
FT 427
FT 428
FT 429
FT 430
FT 431
FT 432
FT 433
FT 434
FT 435
FT 436
FT 437
FT 438
FT 439
FT 440
FT 441
FT 442
FT 443
FT 444
FT 445
FT 446
FT 447
FT 448
FT 449
FT 450
FT 451
FT 452
FT 453
FT 454
FT 455
FT 456
FT 457
FT 458
FT 459
FT 460
FT 461
FT 462
FT 463
FT 464
FT 465
FT 466
FT 467
FT 468
FT 469
FT 470
FT 471
FT 472
FT 473
FT 474
FT 475
FT 476
FT 477
FT 478
FT 479
FT 480
FT 481
FT 482
FT 483
FT 484
FT 485
FT 486
FT 487
FT 488
FT 489
FT 490
FT 491
FT 492
FT 493
FT 494
FT 495
FT 496
FT 497
FT 498
FT 499
FT 500
FT 501
FT 502
FT 503
FT 504
FT 505
FT 506
FT 507
FT 508
FT 509
FT 510
FT 511
FT 512
FT 513
FT 514
FT 515
FT 516
FT 517
FT 518
FT 519
FT 520
FT 521
FT 522
FT 523
FT 524
FT 525
FT 526
FT 527
FT 528
FT 529
FT 530
FT 531
FT 532
FT 533
FT 534
FT 535
FT 536
FT 537
FT 538
FT 539
FT 540
FT 541
FT 542
FT 543
FT 544
FT 545
FT 546
FT 547
FT 548
FT 549
FT 550
FT 551
FT 552
FT 553
FT 554
FT 555
FT 556
FT 557
FT 558
FT 559
FT 560
FT 561
FT 562
FT 563
FT 564
FT 565
FT 566
FT 567
FT 568
FT 569
FT 570
FT 571
FT 572
FT 573
FT 574
FT 575
FT 576
FT 577
FT 578
FT 579
FT 580
FT 581
FT 582
FT 583
FT 584
FT 585
FT 586
FT 587
FT 588
FT 589
FT 590
FT 591
FT 592
FT 593
FT 594
FT 595
FT 596
FT 597
FT 598
FT 599
FT 600
FT 601
FT 602
FT 603
FT 604
FT 605
FT 606
FT 607
FT 608
FT 609
FT 610
FT 611
FT 612
FT 613
FT 614
FT 615
FT 616
FT 617
FT 618
FT 619
FT 620
FT 621
FT 622
FT 623
FT 624
FT 625
FT 626
FT 627
FT 628
FT 629
FT 630
FT 631
FT 632
FT 633
FT 634
FT 635
FT 636
FT 637
FT 638
FT 639
FT 640
FT 641
FT 642
FT 643
FT 644
FT 645
FT 646
FT 647
FT 648
FT 649
FT 650
FT 651
FT 652
FT 653
FT 654
FT 655
FT 656
FT 657
FT 658
FT 659
FT 660
FT 661
FT 662
FT 663
FT 664
FT 665
FT 666
FT 667
FT 668
FT 669
FT 670
FT 671
FT 672
FT 673
FT 674
FT 675
FT 676
FT 677
FT 678
FT 679
FT 680
FT 681
FT 682
FT 683
FT 684
FT 685
FT 686
FT 687
FT 688
FT 689
FT 690
FT 6
```

PT					fragment, Claim 5 *
PP	EP-578616-A.				
PD	12-JAN-1994.				
PD	05-JUL-1993.	810474.			
PD	08-MAR-1993.	AT-000473.			
PR	08-MAY-1993.	AT-000437.			
PR	29-APR-1993.	CH-001310.			
PR	04-MAY-1993.	CH-001375.			
PR	(SANDU) SANDUC LTD.				
PR	(SANDU) SANDUC LTD.				
PR	(SANDU) SANDUC LTD.				
PPI	(SANDU) SANDUC LTD.				
DR	(SANDU) SANDUC LTD.				
DR	wpi: 94-010432/02.				
PT	Isolated DNA sequence - which codes for enzyme having cyclosporin synthetase like activity				
PT	Claim 6; Page 17-41; 93pp; English.				
CC	This sequence encodes an enzyme which has cyclosporin synthetase-				
CC	(formerly known as Tr. initiatum GMS3). The enzyme encoded by this				
CC	sequence catalyses the peptide biosynthesis of cyclosporins and				
CC	structurally related molecules. This sequence may be used for the				
CC	production of cyclosporins or structurally related compounds by				
CC	sequencing in to a recombinant host. This allows effective production				
CC	of antibiotic cyclosporin or its derivatives.				
SQ	Query Match            1.3%    Score 14; DB 9; Length 46899;				
	Best Local Similarity 100.0%; Pred.No. 3.47e+02;				
	Matches     14; Conservative     0; Mismatches     0; Indels     0; Gaps     0;				
Dd	28295 gatccagacggaaa 28308				
Cp	446 gatcgcgaccggaa 433				
ID	VS0287 standard; DNA: 49377 BP.				
AC	V05287:				
DT	21-MAY-1998 (first entry)				
KW	Polyketide synthase from Sorangium cellulosum				
Xw	Sor; biosynthetic module; beta-ketoacyl synthase; acyltransferase;				
KW	keto-reductase; beta-ketone processing domain; cytoskeletal agent;				
KW	intracellular agent; phytopathogenic fungi; transgenic plant;				
OS	Sorangium cellulosum.				
PH	Location/Qualifiers				
FT	Key                383..760				
FT	CD5				/orfcdut5_Sora

```

/note= "gene product highly homologous to the
CDS      927..18974
          /tag= b
          /note= "Saccharopolyspora erythraea"
          /note= "gene product is highly homologous to
          type I PKAs that are known to be involved
          in the synthesis of polyketide compounds"
misc_feature    942..7115
                /tag= c
                /note= "module 1 of SorA"
misc_feature    7203..12884
                /tag= d
                /note= "module 2 of SorA"
misc_feature    13455..15616
                /tag= e
                /note= "module 3 of SorA"
misc_feature    15871..45318
                /tag= f
                /note= "module 4 of SorB"
CDS      19871..46318
          /tag= g
          /product= SorB
          /note= "SorB gene product is highly homologous to
          type I PKS genes"
misc_feature    19870..24556
                /tag= h
                /note= "module 1 of SorB"
misc_feature    24638..30820
                /tag= i
                /note= "module 2 of SorB"
misc_feature    30881..35446
                /tag= j
                /note= "module 3 of SorB"
misc_feature    35528..40214
                /tag= k
                /note= "module 4 of SorB"
misc_feature    40190..46318
                /tag= l
                /note= "module 5 of SorB"
CDS      46851..47891
          /tag= m
          /product= SRN
          /note= "gene product is homologous to the
          methyltransferase from Streptomyces
          hygroscopicus that is involved in
          the synthesis of the polyketide rapamycin"
US5716849-A.
PD PD 10-FEB-1989. 764231
PD PD 14-DEC-1996. 764231
PD PD 2-MAR-1996. 764231

```



[illegible]

RESULT	604
ID	TI8551 standard; DNA: 53577 BP.
AC	TI8551:
DE	06-MAY-1997 (first entry)
DT	Human polycystic kidney disease normal PKD1 gene.
DR	Human polycystic kidney disease; APD; Autosomal dominant;
REV	Multi onset polycystic kidney disease; APD; Autosomal dominant;
REV	mutar. transversion. transition. deletion. insertion. 4

Key	Location/Qualifiers
misc_feature	4379..5772
old_sequence	/note="Specifically claimed region of intronless cDNA identified by exon trapping"
old_sequence	replace(50652..50653, cg)
old_sequence	/tag="c"
old_sequence	/note="Changes Val codon to Leu codon"
old_sequence	replace(50796..50797, cg)
old_sequence	/tag="c"
old_sequence	/note="Changes Val codon by Leu codon"
old_sequence	replace(50827..50828, cc)
old_sequence	/tag="d"
old_sequence	/note="insertion, results in frameshift"

Page 624

**Example 5:** Pages 60-89: 237pp; English.  
**P8** Query Match 100%; Score 14; DB 40; Length 53577;  
**CC** Best Local Similarity 100.0%; Score 14; 3.4e+02;  
**CC** PDB1 focus between the human polycystic kidney disease 1  
**CC** The PKD1 gene or polypeptide may be used to treat autosomal  
**CC** dominant polycystic kidney disease (APKD), and identify carriers  
**CC** of mutant PKD1 genes, i.e. subjects susceptible to APKD. Antibodies  
**CC** also can be used in diagnostic tests. Anti-PKD1 Ab may also be used to  
**CC** perform subcellular and histochemical localization studies, and to  
**CC** block the function of PKD1. AB are also useful in rational drug  
**CC** design studies to identify and test inhibitors of PKD1. Sense and  
**CC** antisense RNA derived from the PKD1 gene may be used for  
**CC** detection and therapy. 8495 A: 17684 C: 15782 G:  
**SQ** Sequence 53577 BP: 8495 A: 17684 C: 15782 G:

---

**P8** Query Match 100%; Score 14; DB 40; Length 53577;  
**CC** Best Local Similarity 100.0%; Score 14; 3.4e+02;  
**CC** PDB1 focus between the human polycystic kidney disease 1  
**CC** The PKD1 gene or polypeptide may be used to treat autosomal  
**CC** dominant polycystic kidney disease (APKD), and identify carriers  
**CC** of mutant PKD1 genes, i.e. subjects susceptible to APKD. Antibodies  
**CC** also can be used in diagnostic tests. Anti-PKD1 Ab may also be used to  
**CC** perform subcellular and histochemical localization studies, and to  
**CC** block the function of PKD1. AB are also useful in rational drug  
**CC** design studies to identify and test inhibitors of PKD1. Sense and  
**CC** antisense RNA derived from the PKD1 gene may be used for  
**CC** detection and therapy. 8495 A: 17684 C: 15782 G:  
**SQ** Sequence 53577 BP: 8495 A: 17684 C: 15782 G:

---

**Db** 19519 tcttgggagcgttt 19532  
**ov** 481 TTTTGGGAGCGCTTT 496

RESULT 606  
 ID NUMBER  
 T1635 standard: DNA: 133894 BP.  
 T1635;  
 AC 03-SEP-1996 (first entry)  
 DT 03-SEP-1996  
 PR 03-SEP-1996  
 CC 03-SEP-1996  
 KW Autophagic nuclear polyhedrosis virus clone 6;  
 KW disruption; non-essential nuclear  
 KW expression vector; baculovirus; ss.  
 PR 03-SEP-1996  
 CC 03-SEP-1996  
 KW Autophagic nuclear polyhedrosis virus clone 6.  
 PR 03-SEP-1996  
 CC 03-SEP-1996  
 DT 30-JUN-1995; B05078.  
 PR 30-JUN-1995; B05078  
 CC 30-JUN-1995; B05078  
 KW (MAY); NATURE; ENVIRONMENT RES COUNCIL.  
 PR Ayres M, Bishop D, Possee R;  
 DT WPI: 96-087670/09.  
 CC 96-087670/09.  
 PR 96-087670/09.  
 CC 96-087670/09.  
 KW Autophagic nuclear polyhedrosis virus complete genome  
 PR sequence - useful in the prodn. of vectors for enhanced  
 DT expression of protein expression, such as interleukin(s),  
 CC interferon(s), and neurotoxin(s).  
 PR Interferon(s), and neurotoxin(s).  
 CC Interferon(s), and neurotoxin(s).  
 KW The complete nucleotide sequence of the genome of clone 6 of the  
 PR baculovirus Autographa californica nuclear polyhedrosis virus (AcNPV)  
 CC has been determined. The sequence is taken from the Genbank record  
 PR J01401.1. The sequence is 127,332 bp long. The sequence is  
 CC open reading frames (ORFs) 11, 20, 22-26, 33-37, 39, 41-46, 50-60.  
 PR open reading frames (ORFs) 11, 20, 22-26, 33-37, 39, 41-46, 50-60.  
 CC open reading frames (ORFs) 11, 20, 22-26, 33-37, 39, 41-46, 50-60.



CC 42-63, 66, 68-70, 81-87, 91-97, 96-98, 101-103, 106-126, 128-130,  
 CC 140-146, 148-150, 152 and 154. Expression vectors contg. the complete  
 CC genomic sequence of AchPV, with the exception that at least one non-  
 CC essential ORF is disrupted or replaced are useful for the synthesis of  
 CC recombinant AchPV.  
 SQ Sequence 13894 BP. 39195 A: 27151 C: 27347 G: 40201 T:  
 Query Match 1.13: Score 14; DB 17; Length 13894;  
 Best Local Similarity 100.0%; Pred. No. 3.47e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 109385 gaactgcgtgactga 109388  
 QY 894 GACTGCCTGCTCA 907  
 |||||  
 RESULT 607  
 ID T10635 standard; DNA; 13894 BP.  
 AC 13894:1-13894 (first entry)  
 DE AcNPV genomic DNA clone 6.  
 KW Autographa californica nuclear polyhedrosis virus clone 6;  
 RW disruption: non-essential gene; heterologous protein production;  
 OS Autographa californica nuclear polyhedrosis virus clone 6.  
 PN N09601320-A2.  
 PD 18-JAN-1996.  
 PF 10-JUN-1995; 180578.  
 PP 10-JUN-1995; 180578.  
 PA Ayes M, Bishop D, Possee R.  
 PI WPI: 96-087670/09.  
 DR GENBANK: L22859.  
 PT sequence - useful in the prodn. of vectors for enhanced  
 PT heterologous protein expression, such as interleukin(s),  
 PT interferon(s) and neurotoxin(s).  
 CC Discourse: Page 90-186; 122pp. English.  
 CC baculovirus Autographa californica nuclear polyhedrosis virus (AcNPV)  
 CC L22859. The patent specification claims a polynucleotide selected from  
 CC 95-53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71,  
 CC 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91,  
 CC 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108,  
 CC 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123,  
 CC 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138,  
 CC 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153,  
 CC 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168,  
 CC 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183,  
 CC 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198,  
 CC 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213,  
 CC 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228,  
 CC 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243,  
 CC 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258,  
 CC 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273,  
 CC 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288,  
 CC 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303,  
 CC 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318,  
 CC 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333,  
 CC 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348,  
 CC 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363,  
 CC 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378,  
 CC 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393,  
 CC 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408,  
 CC 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423,  
 CC 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438,  
 CC 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453,  
 CC 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468,  
 CC 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483,  
 CC 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498,  
 CC 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513,  
 CC 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528,  
 CC 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543,  
 CC 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558,  
 CC 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573,  
 CC 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588,  
 CC 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603,  
 CC 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618,  
 CC 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633,  
 CC 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648,  
 CC 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663,  
 CC 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678,  
 CC 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693,  
 CC 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708,  
 CC 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723,  
 CC 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738,  
 CC 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753,  
 CC 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768,  
 CC 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783,  
 CC 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798,  
 CC 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813,  
 CC 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828,  
 CC 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843,  
 CC 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858,  
 CC 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873,  
 CC 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888,  
 CC 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903,  
 CC 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918,  
 CC 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933,  
 CC 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948,  
 CC 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963,  
 CC 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978,  
 CC 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993,  
 CC 994, 995, 996, 997, 998, 999, 1000, 1001, 1002, 1003, 1004, 1005, 1006, 1007,  
 CC 1008, 1009, 1010, 1011, 1012, 1013, 1014, 1015, 1016, 1017, 1018, 1019, 1020,  
 CC 1021, 1022, 1023, 1024, 1025, 1026, 1027, 1028, 1029, 1030, 1031, 1032, 1033,  
 CC 1034, 1035, 1036, 1037, 1038, 1039, 1040, 1041, 1042, 1043, 1044, 1045, 1046,  
 CC 1047, 1048, 1049, 1050, 1051, 1052, 1053, 1054, 1055, 1056, 1057, 1058,  
 CC 1059, 1060, 1061, 1062, 1063, 1064, 1065, 1066, 1067, 1068, 1069, 1070, 1071,  
 CC 1072, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1080, 1081, 1082, 1083,  
 CC 1084, 1085, 1086, 1087, 1088, 1089, 1090, 1091, 1092, 1093, 1094, 1095, 1096,  
 CC 1097, 1098, 1099, 1100, 1101, 1102, 1103, 1104, 1105, 1106, 1107, 1108, 1109,  
 CC 1110, 1111, 1112, 1113, 1114, 1115, 1116, 1117, 1118, 1119, 1120, 1121, 1122,  
 CC 1123, 1124, 1125, 1126, 1127, 1128, 1129, 1130, 1131, 1132, 1133, 1134, 1135,  
 CC 1136, 1137, 1138, 1139, 1140, 1141, 1142, 1143, 1144, 1145, 1146, 1147, 1148,  
 CC 1149, 1150, 1151, 1152, 1153, 1154, 1155, 1156, 1157, 1158, 1159, 1160, 1161,  
 CC 1162, 1163, 1164, 1165, 1166, 1167, 1168, 1169, 1170, 1171, 1172, 1173, 1174,  
 CC 1175, 1176, 1177, 1178, 1179, 1180, 1181, 1182, 1183, 1184, 1185, 1186, 1187,  
 CC 1188, 1189, 1190, 1191, 1192, 1193, 1194, 1195, 1196, 1197, 1198, 1199, 1200,  
 CC 1201, 1202, 1203, 1204, 1205, 1206, 1207, 1208, 1209, 1210, 1211, 1212, 1213,  
 CC 1214, 1215, 1216, 1217, 1218, 1219, 1220, 1221, 1222, 1223, 1224, 1225, 1226,  
 CC 1227, 1228, 1229, 1230, 1231, 1232, 1233, 1234, 1235, 1236, 1237, 1238, 1239,  
 CC 1240, 1241, 1242, 1243, 1244, 1245, 1246, 1247, 1248, 1249, 1250, 1251, 1252,  
 CC 1253, 1254, 1255, 1256, 1257, 1258, 1259, 1260, 1261, 1262, 1263, 1264, 1265,  
 CC 1266, 1267, 1268, 1269, 1270, 1271, 1272, 1273, 1274, 1275, 1276, 1277, 1278,  
 CC 1279, 1280, 1281, 1282, 1283, 1284, 1285, 1286, 1287, 1288, 1289, 1290, 1291,  
 CC 1292, 1293, 1294, 1295, 1296, 1297, 1298, 1299, 1300, 1301, 1302, 1303, 1304,  
 CC 1305, 1306, 1307, 1308, 1309, 1310, 1311, 1312, 1313, 1314, 1315, 1316, 1317,  
 CC 1318, 1319, 1320, 1321, 1322, 1323, 1324, 1325, 1326, 1327, 1328, 1329, 1330,  
 CC 1331, 1332, 1333, 1334, 1335, 1336, 1337, 1338, 1339, 1340, 1341, 1342, 1343,  
 CC 1344, 1345, 1346, 1347, 1348, 1349, 1350, 1351, 1352, 1353, 1354, 1355, 1356,  
 CC 1357, 1358, 1359, 1360, 1361, 1362, 1363, 1364, 1365, 1366, 1367, 1368, 1369,  
 CC 1370, 1371, 1372, 1373, 1374, 1375, 1376, 1377, 1378, 1379, 1380, 1381, 1382,  
 CC 1383, 1384, 1385, 1386, 1387, 1388, 1389, 1390, 1391, 1392, 1393, 1394, 1395,  
 CC 1396, 1397, 1398, 1399, 1400, 1401, 1402, 1403, 1404, 1405, 1406, 1407, 1408,  
 CC 1409, 1410, 1411, 1412, 1413, 1414, 1415, 1416, 1417, 1418, 1419, 1420, 1421,  
 CC 1422, 1423, 1424, 1425, 1426, 1427, 1428, 1429, 1430, 1431, 1432, 1433, 1434,  
 CC 1435, 1436, 1437, 1438, 1439, 1440, 1441, 1442, 1443, 1444, 1445, 1446, 1447,  
 CC 1448, 1449, 1450, 1451, 1452, 1453, 1454, 1455, 1456, 1457, 1458, 1459, 1460,  
 CC 1461, 1462, 1463, 1464, 1465, 1466, 1467, 1468, 1469, 1470, 1471, 1472, 1473,  
 CC 1474, 1475, 1476, 1477, 1478, 1479, 1480, 1481, 1482, 1483, 1484, 1485, 1486,  
 CC 1487, 1488, 1489, 1490, 1491, 1492, 1493, 1494, 1495, 1496, 1497, 1498, 1499,  
 CC 1500, 1501, 1502, 1503, 1504, 1505, 1506, 1507, 1508, 1509, 1510, 1511, 1512,  
 CC 1513, 1514, 1515, 1516, 1517, 1518, 1519, 1520, 1521, 1522, 1523, 1524, 1525,  
 CC 1526, 1527, 1528, 1529, 1530, 1531, 1532, 1533, 1534, 1535, 1536, 1537, 1538,  
 CC 1539, 1540, 1541, 1542, 1543, 1544, 1545, 1546, 1547, 1548, 1549, 1550, 1551,  
 CC 1552, 1553, 1554, 1555, 1556, 1557, 1558, 1559, 1560, 1561, 1562, 1563, 1564,  
 CC 1565, 1566, 1567, 1568, 1569, 1570, 1571, 1572, 1573, 1574, 1575, 1576, 1577,  
 CC 1578, 1579, 1580, 1581, 1582, 1583, 1584, 1585, 1586, 1587, 1588, 1589, 1590,  
 CC 1591, 1592, 1593, 1594, 1595, 1596, 1597, 1598, 1599, 1600, 1601, 1602, 1603,  
 CC 1604, 1605, 1606, 1607, 1608, 1609, 1610, 1611, 1612, 1613, 1614, 1615, 1616,  
 CC 1617, 1618, 1619, 1620, 1621, 1622, 1623, 1624, 1625, 1626, 1627, 1628, 1629,  
 CC 1630, 1631, 1632, 1633, 1634, 1635, 1636, 1637, 1638, 1639, 1640, 1641, 1642,  
 CC 1643, 1644, 1645, 1646, 1647, 1648, 1649, 1650, 1651, 1652, 1653, 1654, 1655,  
 CC 1656, 1657, 1658, 1659, 1660, 1661, 1662, 1663, 1664, 1665, 1666, 1667, 1668,  
 CC 1669, 1670, 1671, 1672, 1673, 1674, 1675, 1676, 1677, 1678, 1679, 1680, 1681,  
 CC 1682, 1683, 1684, 1685, 1686, 1687, 1688, 1689, 1690, 1691, 1692, 1693, 1694,  
 CC 1695, 1696, 1697, 1698, 1699, 1700, 1701, 1702, 1703, 1704, 1705, 1706, 1707,  
 CC 1708, 1709, 1710, 1711, 1712, 1713, 1714, 1715, 1716, 1717, 1718, 1719, 1720,  
 CC 1721, 1722, 1723, 1724, 1725, 1726, 1727, 1728, 1729, 1730, 1731, 1732, 1733,  
 CC 1734, 1735, 1736, 1737, 1738, 1739, 1740, 1741, 1742, 1743, 1744, 1745, 1746,  
 CC 1747, 1748, 1749, 1750, 1751, 1752, 1753, 1754, 1755, 1756, 1757, 1758, 1759,  
 CC 1760, 1761, 1762, 1763, 1764, 1765, 1766, 1767, 1768, 1769, 1770, 1771, 1772,  
 CC 1773, 1774, 1775, 1776, 1777, 1778, 1779, 1780, 1781, 1782, 1783, 1784, 1785,  
 CC 1786, 1787, 1788, 1789, 1790, 1791, 1792, 1793, 1794, 1795, 1796, 1797, 1798,  
 CC 1799, 1800, 1801, 1802, 1803, 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811,  
 CC 1812, 1813, 1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823, 1824,  
 CC 1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833, 1834, 1835, 1836, 1837,  
 CC 1838, 1839, 1840, 1841, 1842, 1843, 1844, 1845, 1846, 1847, 1848, 1849, 1850,  
 CC 1851, 1852, 1853, 1854, 1855, 1856, 1857, 1858, 1859, 1860, 1861, 1862, 1863,  
 CC 1864, 1865, 1866, 1867, 1868, 1869, 1870, 1871, 1872, 1873, 1874, 1875, 1876,  
 CC 1877, 1878, 1879, 1880, 1881, 1882, 1883, 1884, 1885, 1886, 1887, 1888, 1889,  
 CC 1890, 1891, 1892, 1893, 1894, 1895, 1896, 1897, 1898, 1899, 1900, 1901, 1902,  
 CC 1903, 1904, 1905, 1906, 1907, 1908, 1909, 1910, 1911, 1912, 1913, 1914, 1915,  
 CC 1916, 1917, 1918, 1919, 1920, 1921, 1922, 1923, 1924, 1925, 1926, 1927, 1928,  
 CC 1929, 1930, 1931, 1932, 1933, 1934, 1935, 1936, 1937, 1938, 1939, 1940, 1941,  
 CC 1942, 1943, 1944, 1945, 1946, 1947, 1948, 1949, 1950, 1951, 1952, 1953, 1954,  
 CC 1955, 1956, 1957, 1958, 1959, 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967,  
 CC 1968, 1969, 1970, 1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980,  
 CC 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993,  
 CC 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006,  
 CC 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018



```
PT cds
PT      capsicum*
PT      complement (64988..65731)
PT      /label= MG042
PT      /note= "Previously identified as MORP-19832 and
PT      MORP-20108, the encoded protein shows 41.92
PT      percent identity to the encoded protein from
PT      putrescine transport ATP-binding protein
PT      (potA) from E. coli."
PT      /tag= x
PT      /label= MG043
PT      /note= "Previously identified as MORP-20110, the
PT      encoded protein shows 26.51 percent identity
PT      to spermidine/putrescine transport
PT      system permease protein (potB) from E. coli."
PT      /tag= x
PT      /label= MG044
PT      /note= "Previously identified as MORP-20111, the
PT      encoded protein shows 34.45 percent identity
PT      to spermidine/putrescine transport
PT      system permease protein C (potC) from E. coli."
PT      /tag= x
PT      /label= MG045
PT      /note= "Previously identified as MORP-20112, the
PT      encoded protein shows 36.60 percent identity
PT      to sialoglycoprotease (gcp)
PT      from Pasteurella hemolytica."
PT      /tag= x
PT      /label= MG048
PT      /note= "Previously identified as MORP-19834,
PT      MORP-20109, and MORP-20113, the encoded protein
PT      shows 43.02 percent identity to the encoded
PT      recognition particle protein (rfb) from B.
PT      subtilis."
PT      /tag= x
PT      /label= MG049
PT      /note= "Previously identified as MORP-20114 and
PT      MORP-20115, the encoded protein shows 44.78
PT      percent identity to purine nucleoside
PT      phosphorylase (deob) from E. coli."
PT      /tag= x
PT      /label= MG050
PT      /note= "Previously identified as MORP-20117, the
PT      encoded protein shows 81.03 percent identity
PT      to deoxyribose-phosphate aldolase
```

```
PT cds
PT      (deob) from Mycoplasma pneumoniae"
PT      /tag= x
PT      /label= MG056
PT      /note= "Previously identified as MORP-20122, the
PT      encoded protein shows 30.25 percent
PT      identity to the protein disclosed in
PT      GDI6182_99 from B. subtilis."
PT      /tag= x
PT      /label= MG057
PT      /note= "Previously identified as MORP-20123, the
PT      encoded protein shows 30.60 percent
PT      identity to the protein disclosed in
PT      GDI6185_104 from B. subtilis."
PT      /tag= x
PT      /label= MG067
PT      /note= "Previously identified as MORP-19845, the
PT      encoded protein shows 28.84 percent
PT      identity to glutamic acid specific protease
PT      (gspA) from Staphylococcus aureus."
PT      /tag= x
PT      /label= MG070
PT      /note= "Previously identified as MORP-20136, the
PT      encoded protein shows 28.84 percent
PT      identity to ribonuclease protein S2 (rps2)
PT      from Spiroplasma plantensis."
PT      /tag= x
PT      /label= MG077
PT      /note= "Previously identified as MORP-20140, the
PT      encoded protein shows 28.05 percent
PT      identity to oligopeptide transport system
```

Note: remainder of annotations omitted.

Query Match 1.34; Score 15; DB 27; Length 580073;  
Best Local Similarity 100.04; Pred. No. 8.37e+01;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 578014 cttatcccttgat 578028  
Oy 642 CTTATCCCTTGAT 656

Search completed: Mon Nov 16 15:48:36 1998  
Job time : 294 secs.



!!SEQUENCE LIST 1.0  
(Nucleotide) WORDSEARCH of: /home/obryen/dra977/olig/us-08-887-977-9 check: 9278 from:  
FROMIG of: /home/obryen/dra977/olig/US08887977.seq  
sequence 9, application us/08887977  
general information:  
applicant: wang, wei  
applicant: gish, kurt c.  
applicant: schall, thomas j. . . .

TO: tags: Sequences: 2,336,638 Total-length: 917,356,764 November 16, 1998 17:59

*not 8 sts*  
Database Release Information:

GenBank\_Tags, Release 109.0, Released on 15Oct1998, Formatted on 17Oct1998  
GenBank, Release 109.0, Released on 15Oct1998, Formatted on 17Oct1998  
EMBL\_Tags, Release 56.0, Released on 16Sep1998, Formatted on 18Oct1998  
EMBL, Release 56.0, Released on 16Sep1998, Formatted on 18Oct1998

Word-size: 15 Words: 39166 Diagonals: 2,560 Total-changes: 2,000,000,000  
Integral-width: 1 Alphabet: 4 List-size: 50 CPU minutes: 315.75

Sequence	Strd	Diag	Score	Width	Documentation
GB_EST19:AI045155	-	85	23	1	AI045155 UI-R-CJ-kk-c-08-0-UI-s1 UI-R
GB_EST13:N83058	-	90	20	1	N83058 TgESTzy59g10.r1 TgrH Tachyzoit
GB_EST7:AA255136	-	199	19	1	AA255136 mz82f03.r1 Soares mouse NML
GB_EST7:AA254935	-	-203	19	1	AA254935 mz80h09.r1 Soares mouse NML
GB_EST7:AA245646	-	4	19	1	AA245646 mx01a03.r1 Soares mouse NML
GB_EST6:AA148128	-	-352	19	1	AA148128 z031a06.r1 Stratagene colon
GB_EST5:W84101	-	-682	19	1	W84101 T2925 MVAT4 bloodstream form c
GB_EST5:AA060911	-	-285	19	1	AA060911 mj88b01.r1 Soares mouse p3NM
GB_EST5:AA014588	-	85	19	1	AA014588 mg93a01.r1 Soares mouse embri
GB_GSS2:AQ003850	-	-1013	19	1	AQ003850 CIT-HSP-2288K15.TF CIT-HSP H
GB_EST19:AI047830	+	-152	19	1	AI047830 ud64b12.x1 Sugano mouse live
GB_STS:G39250	+	-35	19	1	G39250 Z20386 Zebrafish AB Danio rerio
GB_EST1:R36281	-	-873	18	1	R36281 y968c11.r1 Homo sapiens cDNA c
GB_EST4:ATTS3711	+	-533	18	1	Z35153 A. thaliana transcribed sequen
GB_EST2:R35058	-	-372	18	1	R95058 yq44f03.r1 Homo sapiens cDNA c
GB_EST2:R83706	-	-354	18	1	R83706 yq14e10.r1 Homo sapiens cDNA c
GB_EST13:AA531233	+	-430	18	1	AA531233 nj52e06.s1 NCI_CGAP_P9 Homc
GB_EST12:AA225058	-	-176	18	1	AA225058 nc21b12.r1 NCI_CGAP_P1 Homc
GB_EST5:W25156	-	-190	18	1	W25156 z869c07.r1 Soares fetal lung N
GB_GSS4:AQ227870	+	-203	18	1	AQ227870 HS_2020.B1_A02_MR CIT Approv
GB_GSS3:AQ224100	-	-136	18	1	AQ224100 HS_2011.A1_E01_MR CIT Approv
GB_GSS3:AQ193807	+	-619	18	1	AQ193807 CIT-HSP-2384O12.TF CIT-HSP H
GB_GSS3:AQ185889	-	-267	18	1	AQ185889 HS_2240.A2_C09_MF CIT Approv
GB_GSS3:AQ133524	-	-133	18	1	AQ133524 HS_3051.A1_B10_MR CIT Approv
GB_GSS3:AQ118433	+	-298	18	1	AQ118433 HS_3007.A2_F10_T7 CIT Approv
GB_GSS1:B35590	-	-129	18	1	B35590 CIT-HSP-2011i1.TR CIT-HSP Homc
GB_GSS1:B31064	-	-690	18	1	B31064 HS_1005-B1-D08-MR.abi CIT Huma
GB_EST20:AI005637	+	-498	18	1	AI005637 cv59b10.s1 Soares testis NHT
GB_EST18:AA919945	+	-789	18	1	AA919945 yv47g04.r1 Stratagene mouse
GB_EST1:T64598	+	102	17	1	T64598 yc25c12.s1 Homo sapiens cDNA c
GB_EST1:T57466	-	-584	17	1	T57466 y856a12.s1 Homo sapiens cDNA c
GB_EST1:T02045	-	-121	17	1	T02045 W8502766 Caenorhabditis elega
GB_EST1:T01726	-	-182	17	1	T01726 W8502447 Caenorhabditis elega
GB_EST5:AA009280	+	-14	17	1	AA009280 mg96c01.r1 Soares mouse embri
GB_EST4:W99065	+	-174	17	1	W99065 mf89g02.r1 Soares mouse embryc
GB_EST4:W98929	+	-629	17	1	W98929 mf91c09.r1 Soares mouse embryc
GB_EST4:W54177	+	45	17	1	W54177 md13a05.r1 Soares mouse embryc
GB_EST4:W44267	+	-170	17	1	W44267 mc74g10.r1 Soares mouse embryc
GB_EST4:W03671	+	-113	17	1	W03671 za65d01.r1 Soares fetal liver
GB_EST4:N79232	+	-551	17	1	N79232 zal2h09.r1 Homo sapiens cDNA c
GB_EST4:W03671	+	-113	17	1	W03671 za65d01.r1 Soares fetal liver
GB_EST4:ATTS4275	+	-135	17	1	Z31244 M.musculus expressed sequence
GB_EST3:N55392	+	-466	17	1	Z45531 A. thaliana transcribed sequen
GB_EST3:HSC10D122	+	-587	17	1	N55392 yz18d11.r1 Homo sapiens cDNA c
GB_EST2:R78075	+	62	17	1	Z41661 H. sapiens partial cDNA sequen
GB_EST2:H02568	+	-829	17	1	R78075 Y180e04.r1 Homo sapiens cDNA c
GB_EST2:H00287	+	-829	17	1	H02568 Y136d05.s1 Homo sapiens cDNA c
GB_EST1:T79571	-	-931	17	1	H00287 Y122h10.r1 Homo sapiens cDNA c
GB_EST13:N81275	-	54	17	1	T79571 yG75h09.r1 Homo sapiens cDNA c
GB_EST13:N69138	-	253	17	1	N81275 TgESTzy25h12.s1 TgrH Tachyzoit
GB_EST13:N69138	-	253	17	1	N69138 TgESTzy33a11.r1 TgrH Tachyzoit

*# of sequences containing matches at least 15 nt long*



Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**CITIZENSHIP**

Result	No.	Score	Query	Match	Length	DB	ID	Description	Prod. No.
c	1	23	7.1	461	19	AI05155		U1-R-C1-Ax-c-08-0-UI-s	2.5ie-14
c	2	20	1.8	412	23	W85077		W85077Human r1 TORH	3.24e-08
c	3	20	1.8	635	22	R38677		50b1 Human retina cDNA	3.24e-08
c	5	19	1.7	276	26	AA245646		mx02.01 r1 Soares mous	2.50e-06
c	6	19	1.7	380	23	W94101		72253 WVAR bloodstrea	2.50e-06
c	7	19	1.7	418	19	W6183		79557777777777777777	2.50e-06
c	8	19	1.7	435	23	AA014588		mg93401.01 Soares mous	2.50e-06
c	10	19	1.7	437	26	AA245935		ms82005.01 Soares mous	2.50e-06
c	11	19	1.7	503	26	AA235318		ms82003.01 Soares mous	2.50e-06
c	12	19	1.7	506	27	AA038850		C1F-SBP-2268K15.TF CIT	2.50e-06
c	13	19	1.7	576	24	AA181487		cs01406.01 Stratagene	2.50e-06
c	15	19	1.7	785	18	AI078930		u64b12.1 r1 Stugano mos	1.50e-04
c	16	18	1.6	317	26	AA275745		u64b12.1 r1 Stratagene	1.50e-04
c	17	18	1.6	377	26	AA275745		U8-1003.H1 D08-MH.apb	1.59e-04
c	18	18	1.6	381	27	B11064		YQ14a10.1 r1 Homo sapien	1.59e-04
c	19	18	1.6	403	20	R37068		YQ14a10.1 r1 Homo sapien	1.59e-04
c	20	18	1.6	435	20	R35558		YQ14a10.1 r1 Homo sapien	1.59e-04
c	21	18	1.6	439	17	AI056307		YQ14a10.1 r1 Homo sapien	1.59e-04
c	22	18	1.6	439	17	AI056307		YQ14a10.1 r1 Homo sapien	1.59e-04
c	23	18	1.6	461	14	R36381		YQ14a10.1 r1 Homo sapien	1.59e-04
c	24	18	1.6	462	23	W25156		YQ14a10.1 r1 Homo sapien	1.59e-04
c	25	18	1.6	491	18	AA250508		nc2121.1 r1 NCI-CPAP.Pr	1.59e-04
c	26	18	1.6	509	8	AA531233		Y152606.01 r1 NCI-CPAP.Pr	1.59e-04
c	28	18	1.6	515	24	R68321		YH58403.01 r1 Homo sapien	1.59e-04
c	29	17	1.5	163	25	AA190203		TH058 HCTCL1 Homo sapi	8.05e-03
c	30	17	1.5	175	6	W0570809		Th058 embryonal carcin	8.05e-03
c	32	17	1.5	215	22	AA451297		Th058 embryonal carcin	8.05e-03
c	33	17	1.5	262	7	AA191212		Th058 embryonal carcin	8.05e-03
c	34	17	1.5	264	8	AA236929		Th058 embryonal carcin	8.05e-03
c	37	17	1.5	283	23	AA00067675		Th058 embryonal carcin	8.05e-03
c	38	17	1.5	286	22	W03271		Th058 embryonal carcin	8.05e-03
c	39	17	1.5	301	18	AI071132		Th058 embryonal carcin	8.05e-03
c	40	17	1.5	302	4	AA336777		ES276496 Cerebellum r1	8.05e-03
c	41	17	1.5	310	8	AA270702		ES276496 Cerebellum r1	8.05e-03
c	42	17	1.5	317	4	AA332497		ES276496 Cerebellum r1	8.05e-03

17	1.5	468 17	AA968418	op49a07.s1	Soares_N7L_	8.05e-0
17	1.5	473 26	AA259539	va48hl2.r1	Soares_mous	8.05e-0

AT032099  
AT063668

9	8	10	17	1	5	481	3	7574662	YB56J42, a1 Homo sapiens	8.05e-03
9	98	17	1	5	482	10	AA018107	YB56J42, a1 Homo sapiens	8.05e-03	
9	100	17	1	5	482	10	AA018107	YB56J42, a1 Homo sapiens	8.05e-03	
9	101	17	1	5	483	18	AI012493	ESL206844, Normalized r	8.05e-03	
9	102	17	1	5	483	18	AI012493	ESL206844, Normalized r	8.05e-03	
9	103	17	1	5	483	23	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	104	17	1	5	484	5	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	105	17	1	5	485	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	106	17	1	5	485	12	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	107	17	1	5	486	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	108	17	1	5	486	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	109	17	1	5	487	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	110	17	1	5	487	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	111	17	1	5	488	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	112	17	1	5	488	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	113	17	1	5	489	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	114	17	1	5	489	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	115	17	1	5	490	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	116	17	1	5	490	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	117	17	1	5	491	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	118	17	1	5	491	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	119	17	1	5	492	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	120	17	1	5	492	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	121	17	1	5	493	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	122	17	1	5	493	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	123	17	1	5	494	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	124	17	1	5	494	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	125	17	1	5	495	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	126	17	1	5	495	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	127	17	1	5	496	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	128	17	1	5	496	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	129	17	1	5	497	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	130	17	1	5	497	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	131	17	1	5	498	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	132	17	1	5	498	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	133	17	1	5	499	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	134	17	1	5	499	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	135	17	1	5	500	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	136	17	1	5	500	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	137	17	1	5	501	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	138	17	1	5	501	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	139	17	1	5	502	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	140	17	1	5	502	18	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	141	17	1	5	503	10	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	142	16	1	4	60	20	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	
9	143	16	1	4	60	20	AA018959	YB56J42, a1 Homo sapiens	8.05e-03	

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

—

```

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668,
```

17	1.5	327 10	AA611184	vo62h02.r1 Soares nous	8.05e-03
17	1.5	336 17	AA963060	UI-R-EI-fv-g-08-0-UI.z	8.05e-03

## 27 AQ0663

48	17	1.5	353	36	AA360146	V47e05-01 Soares monu	8.0e+03
49	17	1.5	353	34	H00287	V12h20-01 Homo sapien	8.0e+03
50	17	1.5	353	34	C50266	C-LEGUS chnua clone y	8.0e+03
51	17	1.5	353	34	C50266	C-LEGUS chnua clone y	8.0e+03
52	17	1.5	370	19	AA061639	TM2713-01 Homo sapien	8.0e+03
53	17	1.5	376	37	B34293	HS-1024-A1-H10-MF, ab	8.0e+03
54	17	1.5	376	37	B34293	HS-1024-A1-H10-MF, ab	8.0e+03
55	17	1.5	377	37	AQ042892	CTT-HSP-23110-TP, CTT	8.0e+03
56	17	1.5	378	9	AA359572	m2h20-02 Homo sapien	8.0e+03
57	17	1.5	378	9	AA359572	m2h20-02 Homo sapien	8.0e+03
58	17	1.5	382	37	B73973	CTT-HSP-2043G10-TP, CTT	8.0e+03
59	17	1.5	385	18	AA028168	ov9e09-01 Soares, test	8.0e+03
60	17	1.5	385	3	764589	YC25012-01 Homo sapien	8.0e+03
61	17	1.5	405	32	AA028168	ov9e09-01 Soares, test	8.0e+03
62	17	1.5	405	32	ATNS74275	m74904-01 Soares monu	8.0e+03
63	17	1.5	408	25	AA316271	m74904-01 Soares monu	8.0e+03
64	17	1.5	413	10	A1028168	m26e03-01 NCI-CGAP, G	8.0e+03
65	17	1.5	414	57	B30574	HS-1003-A1-B09-MR, ab	8.0e+03
66	17	1.5	416	16	AA351068	V79e07-01 Soares monu	8.0e+03
67	17	1.5	417	19	AA060032	UI-8-CL-k4-a-09-0-UI, s	8.0e+03
68	17	1.5	419	9	W64000	TG2E7YF33005-01 T99e4E	8.0e+03
69	17	1.5	419	9	W64000	TG2E7YF33005-01 T99e4E	8.0e+03
70	17	1.5	431	11	AA361382	SW672-01 S45K5K Onchoer	8.0e+03
71	17	1.5	431	18	AA034859	LEu-13p10/70SD Leish	8.0e+03
72	17	1.5	433	12	M65500	20590 Lambda-pRL2 Arab	8.0e+03
73	17	1.5	433	12	MS5332	Y14d11-01 Homo sapien	8.0e+03
74	17	1.5	435	15	AA359572	m2h20-02 Homo sapien	8.0e+03
75	17	1.5	440	15	AA370059	m2h20-02 Homo sapien	8.0e+03
76	17	1.5	442	26	AA361383	m86e09-01 NCI-CGAP, C	8.0e+03
77	17	1.5	443	26	AA361383	m86e09-01 Soares monu	8.0e+03
78	17	1.5	443	26	AA360268	m74e05-01 Soares plene	8.0e+03
79	17	1.5	446	8	AA351107	m92f06-01 NCI-CGAP, C	8.0e+03
80	17	1.5	447	9	AA357001	m3h302-01 Homo sapien	8.0e+03
81	17	1.5	447	9	AA357001	m3h302-01 Homo sapien	8.0e+03
82	17	1.5	448	9	AA357540	n102g05-01 NCI-CGAP, C	8.0e+03
83	17	1.5	449	10	AA318009	v94g05-01 Soares WHH	8.0e+03
84	17	1.5	452	33	H30562	Y18e04-01 Homo sapien	8.0e+03
85	17	1.5	452	33	H30562	Y18e04-01 Homo sapien	8.0e+03
86	17	1.5	455	3	T01726	W620702447 Gaenohabid	8.0e+03
87	17	1.5	458	5	AA214740	v94d05-01 Soares pres	8.0e+03
88	17	1.5	460	44	B36767	YF63d01-01 Homo sapien	8.0e+03
89	17	1.5	460	44	B36767	YF63d01-01 Homo sapien	8.0e+03
90	17	1.5	461	49	AA060045	Y54d11-01 Homo sapien	8.0e+03
91	17	1.5	465	12	AA319843	v94f406-01 Soares plene	8.0e+03
92	17	1.5	467	37	B37271	HS-1023-A1-H08-MF, ab	8.0e+03
93	17	1.5	467	37	B37271	HS-1023-A1-H08-MF, ab	8.0e+03
94	17	1.5	467	37	B37271	Y47H09-01 Homo sapien	8.0e+03







US-08-887-977-9-10t

C	349	16	1.4	428	27	F62071	R. rubripes GSS sequenc	3.20e-01
C	350	16	1.4	428	27	FN07021	T1602737 TAMU Arabidops	3.20e-01
C	351	16	1.4	428	27	R46569	Y55d121 el Homo sapien	3.20e-01
C	352	16	1.4	428	27	R46569	Y55d121 el Homo sapien	3.20e-01
C	353	16	1.4	428	27	AA05331	Y55d121 el Homo sapien	3.20e-01
C	354	16	1.4	433	20	C95941	Y56a01 el Homo sapien	3.20e-01
C	355	16	1.4	433	20	CH9839	Mus musculus early bia	3.20e-01
C	356	16	1.4	432	8	A120881	rs9907 el NCI_GADP_OC	3.20e-01
C	357	16	1.4	432	8	A120881	rs9907 el NCI_GADP_OC	3.20e-01
C	358	16	1.4	432	8	A120881	rs9907 el NCI_GADP_OC	3.20e-01
C	359	16	1.4	437	5	A137127	rs7610 el Stratiagene	3.20e-01
C	360	16	1.4	437	5	A137127	rs7610 el Stratiagene	3.20e-01
C	361	16	1.4	437	5	A137127	rs7610 el Stratiagene	3.20e-01
C	362	16	1.4	437	5	A137127	rs7610 el Stratiagene	3.20e-01
C	363	16	1.4	441	8	AA195491	rs206934 Homo sapien	3.20e-01
C	364	16	1.4	442	6	C18938	rs206934 Homo sapien	3.20e-01
C	365	16	1.4	442	6	C18938	rs206934 Homo sapien	3.20e-01
C	366	16	1.4	442	6	C18938	rs206934 Homo sapien	3.20e-01
C	367	16	1.4	448	6	AA053340	rs206934 Homo sapien	3.20e-01
C	368	16	1.4	448	6	AA053340	rs206934 Homo sapien	3.20e-01
C	369	16	1.4	447	5	AA01942	rs170901 el Soares feta	3.20e-01
C	370	16	1.4	447	5	AA01942	rs170901 el Soares feta	3.20e-01
C	371	16	1.4	447	5	AA01942	rs170901 el Soares feta	3.20e-01
C	372	16	1.4	448	12	AA084316	rs14003 el Stratiagene	3.20e-01
C	373	16	1.4	448	12	AA084316	rs14003 el Stratiagene	3.20e-01
C	374	16	1.4	445	5	AA01958	rs82503 el Soares feta	3.20e-01
C	375	16	1.4	445	5	AA01958	rs82503 el Soares feta	3.20e-01
C	376	16	1.4	451	22	N92505	rs2341 human-PH2 Arab	3.20e-01
C	377	16	1.4	451	22	N92505	rs2341 human-PH2 Arab	3.20e-01
C	378	16	1.4	451	22	N92505	rs2341 human-PH2 Arab	3.20e-01
C	379	16	1.4	451	22	N92505	rs2341 human-PH2 Arab	3.20e-01
C	380	16	1.4	454	27	AA074825	rs2341 human-PH2 Arab	3.20e-01
C	381	16	1.4	454	27	AA074825	rs2341 human-PH2 Arab	3.20e-01
C	382	16	1.4	455	23	AB1087	rs27220605 WATM1 Homo s	3.20e-01
C	383	16	1.4	455	23	AB1087	rs27220605 WATM1 Homo s	3.20e-01
C	384	16	1.4	455	23	AB1087	rs27220605 WATM1 Homo s	3.20e-01
C	385	16	1.4	455	23	AB1087	rs27220605 WATM1 Homo s	3.20e-01
C	386	16	1.4	460	14	RS9341	rs1326207 el Soares monu	3.20e-01
C	387	16	1.4	460	14	RS9341	rs1326207 el Soares monu	3.20e-01
C	388	16	1.4	462	17	R49358	rs1326207 el Soares monu	3.20e-01
C	389	16	1.4	466	6	CO6042	rs2104004 el Stratiagene	3.20e-01
C	390	16	1.4	466	6	CO6042	rs2104004 el Stratiagene	3.20e-01
C	391	16	1.4	467	15	AA188883	rs277611 el Stratiagene	3.20e-01
C	392	16	1.4	467	15	AA188883	rs277611 el Stratiagene	3.20e-01
C	393	16	1.4	468	5	AA370552	rs277611 el Stratiagene	3.20e-01
C	394	16	1.4	468	5	AA370552	rs277611 el Stratiagene	3.20e-01
C	395	16	1.4	469	21	NR20501	rs277611 el Stratiagene	3.20e-01
C	396	16	1.4	469	21	NR20501	rs277611 el Stratiagene	3.20e-01
C	397	16	1.4	470	15	AA055917	rs277611 el Stratiagene	3.20e-01
C	398	16	1.4	470	15	AA055917	rs277611 el Stratiagene	3.20e-01
C	399	16	1.4	471	20	AA748518	rs95908 el NCI_GADP_OC	3.20e-01
C	400	16	1.4	471	20	AA748518	rs95908 el NCI_GADP_OC	3.20e-01
C	401	16	1.4	472	10	R423140	rs95908 el NCI_GADP_OC	3.20e-01
C	402	16	1.4	472	10	R423140	rs95908 el NCI_GADP_OC	3.20e-01
C	403	16	1.4	474	10	AA614935	rs181508 el NCI_GADP_OC	3.20e-01
C	404	16	1.4	474	10	AA614935	rs181508 el NCI_GADP_OC	3.20e-01
C	405	16	1.4	475	25	AA062706	rs770907 el Soas psc	3.20e-01

18-08-887-977-9.rst

C 451	16	1.4	546	23	AA080976	mp68007.r1 Soares mus	3.20e-01
C 452	16	1.4	549	7	AA069574	mp70805.r1 Barstead ms	3.20e-01
C 453	16	1.4	552	5	AA029836	z10c03.r1 Soares pres	3.20e-01
C 454	16	1.4	553	15	AA029836	z10c03.r1 Soares pres	3.20e-01
C 455	16	1.4	554	15	AA173594	sp04902.r1 Stratiopses	3.20e-01
C 456	16	1.4	556	26	AA029836	z10c03.r1 Soares pres	3.20e-01
C 457	16	1.4	557	15	AA029836	z10c03.r1 Soares pres	3.20e-01
C 458	16	1.4	560	12	AA039213	Mus musculus r1 Soares	3.20e-01
C 459	16	1.4	560	12	AA039213	LD1579 Sprime LD Bros	3.20e-01
C 460	16	1.4	562	27	B32172	PC111-2008 799v REX111	3.20e-01
C 461	16	1.4	562	5	C36338	Diervetastium diencioe	3.20e-01
C 462	16	1.4	563	19	AA063322	GS03064.Sprime GH Bros	3.20e-01
C 463	16	1.4	567	27	B32172	PC111-2008 799v REX111	3.20e-01
C 464	16	1.4	567	27	B32172	PC111-2008 799v REX111	3.20e-01
C 465	16	1.4	568	5	AA133920	z10c03.r1 Soares pres	3.20e-01
C 466	16	1.4	572	16	AA079948	ESZ189315 Normalized r	3.20e-01
C 467	16	1.4	574	10	AA060430	79v7802.r1 Knowles Sol	3.20e-01
C 468	16	1.4	574	10	AA060430	79v7802.r1 Knowles Sol	3.20e-01
C 469	16	1.4	577	23	B38575	md7603.r1 Soares mus	3.20e-01
C 470	16	1.4	581	18	C79363	Mus musculus 3.5-dpc b	3.20e-01
C 471	16	1.4	581	21	B56207	CTP-HSP-2006B18 r1r CTR	3.20e-01
C 472	16	1.4	584	16	AA017526	077702.r1 Knowles Sol	3.20e-01
C 473	16	1.4	584	16	AA017526	077702.r1 Knowles Sol	3.20e-01
C 474	16	1.4	592	15	AA0746308	oa542d11.r1 NC1_CGAP_Cc	3.20e-01
C 475	16	1.4	595	7	AA327591	oa542d11.r1 NC1_CGAP_Cc	3.20e-01
C 476	16	1.4	596	12	AA017218	PC111-2008 799v REX111	3.20e-01
C 477	16	1.4	596	12	AA017218	PC111-2008 799v REX111	3.20e-01
C 478	16	1.4	599	13	AA0303177	AA0303177	3.20e-01
C 479	16	1.4	600	27	AA008995	CTP-HSP-2006B18 r1r CTR	3.20e-01
C 480	16	1.4	606	27	AQ041498	CTP-HSP-2006B18 r1r CTR	3.20e-01
C 481	16	1.4	610	15	AA005186	PC111-2008 799v REX111	3.20e-01
C 482	16	1.4	615	5	AA005186	PC111-2008 799v REX111	3.20e-01
C 483	16	1.4	616	24	AA001405	wa45c04.r1 Soares reti	3.20e-01
C 484	16	1.4	617	5	AA005186	PC111-2008 799v REX111	3.20e-01
C 485	16	1.4	618	15	AA005186	PC111-2008 799v REX111	3.20e-01
C 486	16	1.4	618	27	AA005186	PC111-2008 799v REX111	3.20e-01
C 487	16	1.4	619	27	AA0036231	Fugu rubripes GBS sequ	3.20e-01
C 488	16	1.4	619	27	AA0036231	Fugu rubripes GBS sequ	3.20e-01
C 489	16	1.4	621	27	AA004615	077898.r1 Soares NFL	3.20e-01
C 490	16	1.4	621	27	AA004615	077898.r1 Soares NFL	3.20e-01
C 491	16	1.4	621	27	AA0018696	CTP-HSP-2006B18 r1r CTR	3.20e-01
C 492	16	1.4	632	15	AA066611	vz19a03.r1 Soares rN6M	3.20e-01
C 493	16	1.4	636	10	AA061632	9v09f12.r1 Soares mus	3.20e-01
C 494	16	1.4	636	10	AA061632	9v09f12.r1 Soares mus	3.20e-01
C 495	16	1.4	652	17	AA057689	CTP-HSP-2006B18 r1r CTR	3.20e-01
C 496	16	1.4	660	26	AA176612	md7603.r1 Soares mus	3.20e-01
C 497	16	1.4	664	27	B32462	PC111-2008 799v REX111	3.20e-01
C 498	16	1.4	668	20	AA057388	md7603.r1 NC1_CGAP_Cc	3.20e-01
C 499	16	1.4	671	27	AA047818	md7603.r1 NC1_CGAP_Cc	3.20e-01
C 500	16	1.4	671	27	AA047818	md7603.r1 NC1_CGAP_Cc	3.20e-01
C 501	16	1.4	680	25	AA026651	esg6470.r1 Stratiopses	3.20e-01

1.4	476	21	N43231	W78903.r1 Homo sapien	3.20e-01
1.4	477	15	A812668	W45410.r1 Homo sapien	3.20e-01
1.4	478	15	A810409	0588703.r1 NC1_GAAP_GC	3.20e-01
1.4	479	14	F70897	Y149410.r1 Homo sapien	3.20e-01
1.4	480	14	A810598	Y149410.r1 Homo sapien	3.20e-01
1.4	481	14	A810598	Y149410.r1 Homo sapien	3.20e-01
1.4	482	14	F20907	Y4411905.r1 Homo sapien	3.20e-01
1.4	483	14	A8269944	W64411.r1 Soares mous	3.20e-01
1.4	484	14	A8269944	W64411.r1 Soares mous	3.20e-01
1.4	485	14	A8139398	W69511.r1 Sommer Pris	3.20e-01
1.4	486	14	A8139398	W69511.r1 Sommer Pris	3.20e-01
1.4	487	14	A8139398	W69511.r1 Sommer Pris	3.20e-01
1.4	488	14	A8139398	W69511.r1 Sommer Pris	3.20e-01
1.4	489	13	A8145655	WY75810.r1 NC1_GAAP_GC	3.20e-01
1.4	490	13	A8145655	WY75810.r1 NC1_GAAP_GC	3.20e-01
1.4	491	14	A89125	WY75810.r1 NC1_GAAP_GC	3.20e-01
1.4	492	14	A89125	WY75810.r1 NC1_GAAP_GC	3.20e-01
1.4	493	14	A8134162	WY14908.r1 Soares test	3.20e-01
1.4	494	14	A8134162	WY14908.r1 Soares test	3.20e-01
1.4	495	14	A8134162	WY14908.r1 Soares test	3.20e-01
1.4	496	14	A8134162	WY14908.r1 Soares test	3.20e-01
1.4	497	14	A8134162	WY14908.r1 Soares test	3.20e-01
1.4	498	14	A8134162	WY14908.r1 Soares test	3.20e-01
1.4	499	14	A8134162	WY14908.r1 Soares test	3.20e-01
1.4	500	14	A8137056	WY14908.r1 Soares test	3.20e-01
1.4	501	3	W75905	CE20635 Cenorhobabitis	3.20e-01
1.4	502	14	A8142007	WY67901.r1 Sugamo mous	3.20e-01
1.4	503	14	A8142007	WY67901.r1 Sugamo mous	3.20e-01
1.4	504	14	A8142007	WY67901.r1 Sugamo mous	3.20e-01
1.4	505	14	A8142007	WY67901.r1 Sugamo mous	3.20e-01
1.4	506	14	A8142007	WY67901.r1 Sugamo mous	3.20e-01
1.4	507	14	A8141168	WY67901.r1 Sugamo mous	3.20e-01
1.4	508	14	A8141168	WY67901.r1 Sugamo mous	3.20e-01
1.4	509	14	A8141168	WY67901.r1 Sugamo mous	3.20e-01
1.4	510	14	A8141168	WY67901.r1 Sugamo mous	3.20e-01
1.4	511	17	A8595910	Y454951.r1 Soares ovar	3.20e-01
1.4	512	17	A8595910	Y454951.r1 Soares ovar	3.20e-01
1.4	513	17	A8595910	Y454951.r1 Soares ovar	3.20e-01
1.4	514	17	A8595910	Y454951.r1 Soares ovar	3.20e-01
1.4	515	22	W73544	Y454951.r1 Soares ovar	3.20e-01
1.4	516	22	W73544	Y454951.r1 Soares ovar	3.20e-01
1.4	517	22	W73544	Y454951.r1 Soares ovar	3.20e-01
1.4	518	22	W73544	Y454951.r1 Soares ovar	3.20e-01
1.4	519	5	A8387279	Y454951.r1 Soares ovar	3.20e-01
1.4	520	17	A8419048	Y454951.r1 Soares ovar	3.20e-01
1.4	521	17	A8419048	Y454951.r1 Soares ovar	3.20e-01
1.4	522	16	C80803	Y454951.r1 Soares ovar	3.20e-01
1.4	523	16	C80803	Y454951.r1 Soares ovar	3.20e-01
1.4	524	24	A8100927	GM41200 Sprime GM Dros	3.20e-01
1.4	525	24	A8100927	GM41200 Sprime GM Dros	3.20e-01
1.4	526	24	A8100927	GM41200 Sprime GM Dros	3.20e-01
1.4	527	11	A8424244	Y454951.r1 Soares para	3.20e-01
1.4	528	11	A8424244	Y454951.r1 Soares para	3.20e-01
1.4	529	11	A8424244	Y454951.r1 Soares para	3.20e-01
1.4	530	11	A8424244	Y454951.r1 Soares para	3.20e-01
1.4	531	13	A8802463	GM55396 Sprime GM Dros	3.20e-01
1.4	532	25	A8802463	GM55396 Sprime GM Dros	3.20e-01
1.4	533	25	A8802463	GM55396 Sprime GM Dros	3.20e-01
1.4	534	13	A8803728	GM41200 Sprime GM Dros	3.20e-01

608	13	MA461112	E330604, Sprime LD Dros	1.30e-01
609	13	MA461112	E330604, Sprime LD Dros	1.30e-01
610	13	MA461112	E330604, Sprime LD Dros	1.30e-01
611	14	702118	CT18570-2051324, TP C1T	3.30e-01
612	14	702118	CT18570-2051324, TP C1T	3.30e-01
613	14	722127	ab50503, al Saguano mouse	3.30e-01
614	14	722128	ab50503, al Saguano mouse	3.30e-01
615	14	731205	g530604, al Strateges	3.30e-01
616	14	731205	g530604, al Strateges	3.30e-01
617	14	805119	g530604, al Strateges	3.30e-01
618	14	805119	g530604, al Strateges	3.30e-01
619	14	805119	g530604, al Strateges	3.30e-01
620	14	805119	g530604, al Strateges	3.30e-01
621	14	805119	g530604, al Strateges	3.30e-01
622	14	805119	g530604, al Strateges	3.30e-01
623	14	805119	g530604, al Strateges	3.30e-01
624	14	805119	g530604, al Strateges	3.30e-01
625	14	805119	g530604, al Strateges	3.30e-01
626	14	805119	g530604, al Strateges	3.30e-01
627	14	805119	g530604, al Strateges	3.30e-01
628	14	805119	g530604, al Strateges	3.30e-01
629	14	805119	g530604, al Strateges	3.30e-01
630	14	805119	g530604, al Strateges	3.30e-01
631	14	805119	g530604, al Strateges	3.30e-01
632	14	805119	g530604, al Strateges	3.30e-01
633	14	805119	g530604, al Strateges	3.30e-01
634	14	805119	g530604, al Strateges	3.30e-01
635	14	805119	g530604, al Strateges	3.30e-01
636	14	805119	g530604, al Strateges	3.30e-01
637	14	805119	g530604, al Strateges	3.30e-01
638	14	805119	g530604, al Strateges	3.30e-01
639	14	805119	g530604, al Strateges	3.30e-01
640	14	805119	g530604, al Strateges	3.30e-01
641	14	805119	g530604, al Strateges	3.30e-01
642	14	805119	g530604, al Strateges	3.30e-01
643	14	805119	g530604, al Strateges	3.30e-01
644	14	805119	g530604, al Strateges	3.30e-01
645	14	805119	g530604, al Strateges	3.30e-01
646	14	805119	g530604, al Strateges	3.30e-01
647	14	805119	g530604, al Strateges	3.30e-01
648	14	805119	g530604, al Strateges	3.30e-01
649	14	805119	g530604, al Strateges	3.30e-01
650	14	805119	g530604, al Strateges	3.30e-01
651	14	805119	g530604, al Strateges	3.30e-01
652	14	805119	g530604, al Strateges	3.30e-01
653	14	805119	g530604, al Strateges	3.30e-01
654	14	805119	g530604, al Strateges	3.30e-01
655	14	805119	g530604, al Strateges	3.30e-01
656	14	805119	g530604, al Strateges	3.30e-01
657	14	805119	g530604, al Strateges	3.30e-01
658	14	805119	g530604, al Strateges	3.30e-01
659	14	805119	g530604, al Strateges	3.30e-01
660	14	805119	g530604, al Strateges	3.30e-01
661	14	805119	g530604, al Strateges	3.30e-01
662	14	805119	g530604, al Strateges	3.30e-01
663	14	805119	g530604, al Strateges	3.30e-01
664	14	805119	g530604, al Strateges	3.30e-01
665	14	805119	g530604, al Strateges	3.30e-01
666	14	805119	g530604, al Strateges	3.30e-01
667	14	805119	g530604, al Strateges	3.30e-01
668	14	805119	g530604, al Strateges	3.30e-01
669	14	805119	g530604, al Strateges	3.30e-01
670	14	805119	g530604, al Strateges	3.30e-01
671	14	805119	g530604, al Strateges	3.30e-01
672	14	805119	g530604, al Strateges	3.30e-01
673	14	805119	g530604, al Strateges	3.30e-01
674	14	805119	g530604, al Strateges	3.30e-01
675	14	805119	g530604, al Strateges	3.30e-01
676	14	805119	g530604, al Strateges	3.30e-01
677	14	805119	g530604, al Strateges	3.30e-01
678	14	805119	g530604, al Strateges	3.30e-01
679	14	805119	g530604, al Strateges	3.30e-01
680	14	805119	g530604, al Strateges	3.30e-01
681	14	805119	g530604, al Strateges	3.30e-01
682	14	805119	g530604, al Strateges	3.30e-01
683	14	805119	g530604, al Strateges	3.30e-01
684	14	805119	g530604, al Strateges	3.30e-01
685	14	805119	g530604, al Strateges	3.30e-01
686	14	805119	g530604, al Strateges	3.30e-01
687	14	805119	g530604, al Strateges	3.30e-01
688	14	805119	g530604, al Strateges	3.30e-01
689	14	805119	g530604, al Strateges	3.30e-01
690	14	805119	g530604, al Strateges	3.30e-01
691	14	805119	g530604, al Strateges	3.30e-01
692	14	805119	g530604, al Strateges	3.30e-01
693	14	805119	g530604, al Strateges	3.30e-01
694	14	805119	g530604, al Strateges	3.30e-01
695	14	805119	g530604, al Strateges	3.30e-01
696	14	805119	g530604, al Strateges	3.30e-01
697	14	805119	g530604, al Strateges	3.30e-01
698	14	805119	g530604, al Strateges	3.30e-01
699	14	805119	g530604, al Strateges	3.30e-01
700	14	805119	g530604, al Strateges	3.30e-01
701	14	805119	g530604, al Strateges	3.30e-01
702	14	805119	g530604, al Strateges	3.30e-01
703	14	805119	g530604, al Strateges	3.30e-01
704	14	805119	g530604, al Strateges	3.30e-01
705	14	805119	g530604, al Strateges	3.30e-01
706	14	805119	g530604, al Strateges	3.30e-01
707	14	805119	g530604, al Strateges	3.30e-01
708	14	805119	g530604, al Strateges	3.30e-01
709	14	805119	g530604, al Strateges	3.30e-01
710	14	805119	g530604, al Strateges	3.30e-01
711	14	805119	g530604, al Strateges	3.30e-01
712	14	805119	g530604, al Strateges	3.30e-01
713	14	805119	g530604, al Strateges	3.30e-01
714	14	805119	g530604, al Strateges	3.30e-01
715	14	805119	g530604, al Strateges	3.30e-01
716	14	805119	g530604, al Strateges	3.30e-01
717	14	805119	g530604, al Strateges	3.30e-01
718	14	805119	g530604, al Strateges	3.30e-01
719	14	805119	g530604, al Strateges	3.30e-01
720	14	805119	g530604, al Strateges	3.30e-01
721	14	805119	g530604, al Strateges	3.30e-01
722	14	805119	g530604, al Strateges	3.30e-01
723	14	805119	g530604, al Strateges	3.30e-01
724	14	805119	g530604, al Strateges	3.30e-01
725	14	805119	g530604, al Strateges	3.30e-01
726	14	805119	g530604, al Strateges	3.30e-01
727	14	805119	g530604, al Strateges	3.30e-01
728	14	805119	g530604, al Strateges	3.30e-01
729	14	805119	g530604, al Strateges	3.30e-01
730	14	805119	g530604, al Strateges	3.30e-01
731	14	805119	g530604, al Strateges	3.30e-01
732	14	805119	g530604, al Strateges	3.30e-01
733	14	805119	g530604, al Strateges	3.30e-01
734	14	805119	g530604, al Strateges	3.30e-01
735	14	805119	g530604, al Strateges	3.30e-01
736	14	805119	g530604, al Strateges	3.30e-01
737	14	805119	g530604, al Strateges	3.30e-01
738	14	805119	g530604, al Strateges	3.30e-01
739	14	805119	g530604, al Strateges	3.30e-01
740	14	805119	g530604, al Strateges	3.30e-01
741	14	805119	g530604, al Strateges	3.30e-01
742	14	805119	g530604, al Strateges	3.30e-01
743	14	805119	g530604, al Strateges	3.30e-01
744	14	805119	g530604, al Strateges	3.30e-01
745	14	805119	g530604, al Strateges	3.30e-01
746	14	805119	g530604, al Strateges	3.30e-01
747	14	805119	g530604, al Strateges	3.30e-01
748	14	805119	g530604, al Strateges	3.30e-01
749	14	805119	g530604, al Strateges	3.30e-01
750	14	805119	g530604, al Strateges	3.30e-01
751	14	805119	g530604, al Strateges	3.30e-01
752	14	805119	g530604, al Strateges	3.30e-01
753	14	805119	g530604, al Strateges	3.30e-01
754	14	805119	g530604, al Strateges	3.30e-01
755	14	805119	g530604, al Strateges	3.30e-01
756	14	805119	g530604, al Strateges	3.30e-01
757	14	805119	g530604, al Strateges	3.30e-01
758	14	805119	g530604, al Strateges	3.30e-01
759	14	805119	g530604, al Strateges	3.30e-01
760	14	805119	g530604, al Strateges	3.30e-01
761	14	805119	g530604, al Strateges	3.30e-01
762	14	805119	g530604, al Strateges	3.30e-01
763	14	805119	g530604, al Strateges	3.30e-01
764	14	805119	g530604, al Strateges	3.30e-01
765	14	805119	g530604, al Strateges	3.30e-01
766	14	805119	g530604, al Strateges	3.30e-01
767	14	805119	g530604, al Strateges	3.30e-01
768	14	805119	g530604, al Strateges	3.30e-01
769	14	805119	g530604, al Strateges	3.30e-01
770	14	805119	g530604, al Strateges	3.30e-01
771	14	805119	g530604, al Strateges	3.30e-01
772	14	805119	g530604, al Strateges	3.30e-01
773	14	805119	g530604, al Strateges	3.30e-01
774	14	805119	g530604, al Strateges	3.30e-01
775	14	805119	g530604, al Strateges	3.30e-01
776	14	805119	g530604, al Strateges	3.30e-01
777	14	805119	g530604, al Strateges	3.30e-01
778	14	805119	g530604, al Strateges	3.30e-01
779	14	805119	g530604, al Strateges	3.30e-01
780	14	805119	g530604, al Strateges	3.30e-01
781	14	805119	g530604, al Strateges	3.30e-01
782	14	805119	g530604, al Strateges	3.30e-01
783	14	805119	g530604, al Strateges	3.30e-01
784	14	805119	g530604, al Strateges	3.30e-01
785	14	805119	g530604, al Strateges	3.30e-01
786	14	805119	g530604, al Strateges	3.30e-01
787	14	805119	g530604, al Strateges	3.30e-01
788	14	805119	g530604, al Strateges	3.30e-01
789	14	805119	g530604, al Strateges	3.30e-01
790	14	805119	g530604, al Strateges	3.30e-01
791	14	805119	g530604, al Strateges	3.30e-01
792	14	805119	g530604, al Strateges	3.30e-01
793	14	805119	g530604, al Strateges	3.30e-01
794	14	805119	g530604, al Strateges	3.30e-01
795	14	805119	g530604, al Strateges	3.30e-01
796	14	805119	g530604, al Strateges	3.30e-01
797	14	805119	g530604, al Strateges	3.30e-01
798	14	805119	g530604, al Strateges	3.30e-01
799	14	805119	g530604, al Strateges	3.30e-01
800	14	805119	g530604, al Strateges	3.30e-01







c	757	15	1.3	43	5	AA03782	rs49102 r1 Soares preg	9.65e+00
c	758	15	1.3	434	9	AA57030	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	759	15	1.3	434	9	AA57030	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	760	15	1.3	434	17	AA57230	rs49102 r1 Stratagene	9.65e+00
c	761	15	1.3	436	8	AA57030	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	762	15	1.3	436	8	AA57030	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	763	15	1.3	437	11	AA57394	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	764	15	1.3	437	27	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	765	15	1.3	437	27	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	766	15	1.3	438	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	767	15	1.3	438	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	768	15	1.3	439	22	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	769	15	1.3	440	6	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	770	15	1.3	440	6	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	771	15	1.3	440	6	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	772	15	1.3	441	21	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	773	15	1.3	441	21	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	774	15	1.3	442	24	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	775	15	1.3	442	24	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	776	15	1.3	442	24	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	777	15	1.3	443	24	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	778	15	1.3	443	24	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	779	15	1.3	444	21	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	780	15	1.3	444	21	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	781	15	1.3	445	24	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	782	15	1.3	445	24	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	783	15	1.3	446	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	784	15	1.3	446	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	785	15	1.3	447	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	786	15	1.3	448	5	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	787	15	1.3	448	5	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	788	15	1.3	448	5	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	789	15	1.3	449	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	790	15	1.3	449	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	791	15	1.3	450	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	792	15	1.3	450	16	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	793	15	1.3	451	22	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	794	15	1.3	451	22	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	795	15	1.3	452	20	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	796	15	1.3	453	18	AA046549	rs49102 r1 Soares preg	9.65e+00
c	797	15	1.3	453	17	AA042420	rs49102 r1 Soares preg	9.65e+00
c	798	15	1.3	454	14	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	799	15	1.3	454	14	AA043040	rs179303 r1 Barstead mo	9.65e+00
c	800	15	1.3	455	20	AA046549	rs49102 r1 Soares preg	9.65e+00
c	801	15	1.3	456	16	AA047423	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	802	15	1.3	456	16	AA047423	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	803	15	1.3	457	15	AA045450	rs49102 r1 Soares preg	9.65e+00
c	804	15	1.3	457	15	AA045450	rs49102 r1 Soares preg	9.65e+00
c	805	15	1.3	458	20	AA046549	rs49102 r1 Soares preg	9.65e+00
c	806	15	1.3	459	18	AA047096	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	807	15	1.3	460	15	AA047096	rs49102 r1 NCI_CGAP_P2	9.65e+00
c	808	15	1.3	461	15	AA045450	rs49102 r1 Soares preg	9.65e+00
c	809	15	1.3	461	15	AA045450	rs49102 r1 Soares preg	9.65e+00

C 808	15	1.3	461.18	A0170262	wd35901.1l Soares_sense	9.55±00
C 809	15	1.3	461.25	A0170261	wd35903.1l Soares_sense	9.55±00
C 810	15	1.3	461.29	A0170382	wd37203.1l Soares_NSE	9.55±00
C 811	15	1.3	461.33	W67529	wd4108.1l Soares_feta	9.55±00
C 812	15	1.3	461.35	W67529	wd4108.1l Soares_feta	9.55±00
C 813	15	1.3	462.25	AL104342	mg85905.1l Stratyagene	9.55±00
C 814	15	1.3	463.11	AL681467	vt37901.1l Barteadm	9.55±00
C 815	15	1.3	463.26	AA256504	mx01e2.1l Soares_mous	9.55±00
C 816	15	1.3	463.32	AA256504	mx01e2.1l Soares_mous	9.55±00
C 817	15	1.3	463.35	AA256504	vt19704.1l NCLCGAP.Pr	9.55±00
C 818	15	1.3	463.42	AA256504	vt19704.1l NCLCGAP.Pr	9.55±00
C 819	15	1.3	464.6	TC93493A	Rice_CDM, partial seq	9.55±00
C 820	15	1.3	465.3	W53196	yl93902.1l Homo sapien	9.55±00
C 821	15	1.3	466.5	AL396872	yl93902.1l Life Tech	9.55±00
C 822	15	1.3	466.21	W7737147	yl93910.1l Homo sapien	9.55±00
C 823	15	1.3	467.26	AL14766	mr7367.1l Soares_mous	9.55±00
C 824	15	1.3	467.26	AL163711	mr22609.1l Soares_mous	9.55±00
C 825	15	1.3	467.31	AL163711	vt37902.1l Soares_test	9.55±00
C 826	15	1.3	467.31	AL163711	vt37902.1l Soares_test	9.55±00
C 827	15	1.3	468.5	AL11006	wd30164.1l Soares_WHM	9.55±00
C 828	15	1.3	469.13	AL71007	vt95906.6l NCLCGAP_CC	9.55±00
C 829	15	1.3	471.9	W53196	tg25929yghml.1l TgE494	9.55±00
C 830	15	1.3	471.9	W53196	tg25929yghml.1l TgE494	9.55±00
C 831	15	1.3	471.27	AL461151	wd46802.1l Soares_feta	9.55±00
C 832	15	1.3	473.7	AL814611	wb97467.1l NCLCGAP_CC	9.55±00
C 833	15	1.3	473.13	AL771334	ac262e2.1l Stratyagene	9.55±00
C 834	15	1.3	474.21	AL771334	yl74407.1l Homo sapien	9.55±00
C 835	15	1.3	474.21	AL771334	yl74407.1l Homo sapien	9.55±00
C 836	15	1.3	476.16	AL877711	nr09801.1l NCLCGAP_CO	9.55±00
C 837	15	1.3	476.13	AL755758	vt7431.1l Stratyagene	9.55±00
C 838	15	1.3	477.7	AL435249	wd43901.1l Knowles_sol	9.55±00
C 839	15	1.3	477.7	AL435249	wd43901.1l Knowles_sol	9.55±00
C 840	15	1.3	477.10	AL265050	wd49601.1l Soares_tota	9.55±00
C 841	15	1.3	477.19	A011847	Schizosaccharomyces pos	9.55±00
C 842	15	1.3	478.15	AL635394	of28210.1l Soares_test	9.55±00
C 843	15	1.3	478.15	AL635394	of28210.1l Soares_test	9.55±00
C 844	15	1.3	478.13	AL398962	mg01004.1l Soares_prey	9.55±00
C 845	15	1.3	479.26	AL139566	mg01004.1l Soares_prey	9.55±00
C 846	15	1.3	480.20	H30852	vt708212.1l Homo sapien	9.55±00
C 847	15	1.3	481.15	AL58947	UT-R-A0-aj3-b-02-U-1a	9.55±00
C 848	15	1.3	481.15	AL58947	UT-R-A0-aj3-b-02-U-1a	9.55±00
C 849	15	1.3	482.5	A035067	mg13409.1l Soares_WHM	9.55±00
C 850	15	1.3	482.7	AL493558	vl83409.1l Stratyagene	9.55±00
C 851	15	1.3	482.7	A0003936	CT-HSP-22H59117.CIT	9.55±00
C 852	15	1.3	484.10	AL58947	MG004-NEW_Canada_fap	9.55±00
C 853	15	1.3	484.10	AL58947	MG004-NEW_Canada_fap	9.55±00
C 854	15	1.3	484.24	AL61604	vt54606.1l Knowles_sol	9.55±00
C 855	15	1.3	484.13	W16342	zc16c05.1l Soares_para	9.55±00
C 856	15	1.3	485.17	AL713257	vt97511.1l Soares_NFL	9.55±00
C 857	15	1.3	485.17	AL713257	vt97511.1l Soares_NFL	9.55±00
C 858	15	1.3	486.8	AA563134	vl81805.1l NCLCGAP_Pr	9.55±00

850	15	1,3	487	6	C17732	Human placenta cDNA 5'	9.65e+00
851	15	1,3	489	25	A4217307	mu2n02.r1 Soares moss	9.65e+00
852	15	1,3	490	23	T66590	50510.12 Soares moss	9.65e+00
853	15	1,3	491	23	T66590	50510.12 Soares moss	9.65e+00
854	15	1,3	491	1	A4364065	51510.11 Soares moss	9.65e+00
845	15	1,3	491	18	A1077503	as4912.1 Soares rHM	9.65e+00
855	15	1,3	491	18	A1077503	ESR199598 Normalized r	9.65e+00
856	15	1,3	494	27	A4217307	140c12.11 Biddington	9.65e+00
857	15	1,3	494	27	A4217307	140c12.11 Biddington	9.65e+00
858	15	1,3	495	5	A4113338	1929 f09.r1 Soares moss	9.65e+00
859	15	1,3	495	5	A4113338	1929 f09.r1 Soares moss	9.65e+00
860	15	1,3	496	9	AA586201	28850 Lambda-PHL2.L	9.65e+00
861	15	1,3	496	27	B55571	CIT-ESP-39701.r1 CT CIP-	9.65e+00
862	15	1,3	500	1	A4364065	51510.11 Soares moss	9.65e+00
863	15	1,3	501	26	A4364065	51510.11 Soares moss	9.65e+00
864	15	1,3	501	26	AA086646	601311.r1 Stratagene	9.65e+00
865	15	1,3	501	18	AA933796	957408.r1 Soares rHM	9.65e+00
866	15	1,3	501	18	AA933796	957408.r1 Soares rHM	9.65e+00
867	15	1,3	501	6	AA933796	957408.r1 Soares rHM	9.65e+00
868	15	1,3	501	6	AA933796	957408.r1 Soares rHM	9.65e+00
869	15	1,3	501	6	AA933796	957408.r1 Soares rHM	9.65e+00
870	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
871	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
872	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
873	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
874	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
875	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
876	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
877	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
878	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
879	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
880	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
881	15	1,3	504	21	AA607662	mm1310.r1 Stratagene	9.65e+00
882	15	1,3	511	15	AA450508	407602.r1 Knowles Sol	9.65e+00
883	15	1,3	511	15	AA450508	407602.r1 Knowles Sol	9.65e+00
884	15	1,3	512	10	AA456933	2024007.r1 Soares rHM	9.65e+00
885	15	1,3	512	10	AA456933	2024007.r1 Soares rHM	9.65e+00
886	15	1,3	512	5	AA555062	418506.r1 Soares rHM	9.65e+00
887	15	1,3	512	5	AA555062	418506.r1 Soares rHM	9.65e+00
888	15	1,3	512	5	AA586804	mm59c06.r1 Stratagene	9.65e+00
889	15	1,3	512	9	AA586804	mm59c06.r1 Stratagene	9.65e+00
890	15	1,3	512	9	AA586804	mm59c06.r1 Stratagene	9.65e+00
891	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
892	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
893	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
894	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
895	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
896	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
897	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
898	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
899	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
900	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
901	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
902	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
903	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
904	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
905	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
906	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
907	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
908	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
909	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00
910	15	1,3	512	17	AA586804	mm59c06.r1 Stratagene	9.65e+00

c	910	15	1.3	547	A5177117	v93p11 v1 Knowles Sol	9.55+00
c	911	15	1.3	548	A5177141	mA3c007 v1 Life Tech	9.55+00
c	912	15	1.3	551	A514661	mm0a02 v1 Reddington	9.55+00
c	913	15	1.3	551	A514661	v6c004 v1 Ko mussion	9.55+00
c	914	15	1.3	553	A5197428	mm0a02 v1 Reddington	9.55+00
c	915	15	1.3	554	A5155145	mm8b06 v1 Stratageme	9.55+00
c	916	15	1.3	554	A5155145	mm8b06 v1 Stratageme	9.55+00
c	917	15	1.3	554	A5155145	A. Chailiana transcribe	9.55+00
c	918	15	1.3	556	A5221967	mm5d10 v1 Soares mus	9.55+00
c	919	15	1.3	556	A5221967	mm5d10 v1 Soares mus	9.55+00
c	920	15	1.3	557	B90353	mm5d10 v1 Soares mus	9.55+00
c	921	15	1.3	561	A5189924	mm5d10 v1 Soares mus	9.55+00
c	922	15	1.3	563	A5190372	v42f5b02 v1 Baratead	9.55+00
c	923	15	1.3	563	A5190372	v42f5b02 v1 Baratead	9.55+00
c	924	15	1.3	566	A9003775	v41a1003 v1 Soares mus	9.55+00
c	925	15	1.3	569	A9003775	v41a1003 v1 Soares mus	9.55+00
c	926	15	1.3	569	A9037050	mm2f012 v1 Stratageme	9.55+00
c	927	15	1.3	570	A9045976	mm2f012 v1 Stratageme	9.55+00
c	928	15	1.3	572	A5146670	v072a002 v1 Knowles Sol	9.55+00
c	929	15	1.3	572	A5146670	v072a002 v1 Knowles Sol	9.55+00
c	930	15	1.3	574	A5194238	C17-HSP-702N141/188 C17	9.55+00
c	931	15	1.3	574	A5194238	vA74003 v1 Knowles Sol	9.55+00
c	932	15	1.3	575	A5392754	C17-HSP-100A17/78B C17	9.55+00
c	933	15	1.3	575	A5392754	C17-HSP-100A17/78B C17	9.55+00
c	934	15	1.3	577	B90497	C17-HSP-100A17/78B C17	9.55+00
c	935	15	1.3	577	A519477	nn78f02 v1 Stratageme	9.55+00
c	936	15	1.3	578	A5203561	mm5d02 v1 Soares feta	9.55+00
c	937	15	1.3	578	A5203561	mm5d02 v1 Soares feta	9.55+00
c	938	15	1.3	581	A5198585	mm5d02 v1 Soares feta	9.55+00
c	939	15	1.3	585	A5042762	v533h06 v1 Homo sapien	9.55+00
c	940	15	1.3	585	A5042762	v533h06 v1 Homo sapien	9.55+00
c	941	15	1.3	587	A1009516	mm5f03 v1 Stratageme	9.55+00
c	942	15	1.3	590	A1671535	mm5f03 v1 Stratageme	9.55+00
c	943	15	1.3	591	A1042823	v55f212 v1 Sugano mus	9.55+00
c	944	15	1.3	591	A1042823	v55f212 v1 Sugano mus	9.55+00
c	945	15	1.3	591	A5200854	Tg25f25x4h03 v1 TgH2549	9.55+00
c	946	15	1.3	591	A5200854	Tg25f25x4h03 v1 TgH2549	9.55+00
c	947	15	1.3	596	A518970	T5622TFR TMM Barbidop	9.55+00
c	948	15	1.3	596	A518970	v109a06 v1 Barbidop	9.55+00
c	949	15	1.3	601	L87919	mm5d03 v1 3-dpc b	9.55+00
c	950	15	1.3	602	A5172078	mm5d03 v1 NCI_C6p1	9.55+00
c	951	15	1.3	605	A5172078	mm5d03 v1 NCI_C6p1	9.55+00
c	952	15	1.3	607	A5176882	mm5d03 v1 NCI_C6p1	9.55+00
c	953	15	1.3	611	A5911686	mm5d03 v1 NCI_C6p1	9.55+00
c	954	15	1.3	611	A5911686	mm5d03 v1 NCI_C6p1	9.55+00
c	955	15	1.3	615	A5165735	mm5d03 v1 NCI_C6p1	9.55+00
c	956	15	1.3	620	PR0004712	mm5d03 v1 NCI_C6p1	9.55+00
c	957	15	1.3	620	A5042519	mm5d03 v1 NCI_C6p1	9.55+00
c	958	15	1.3	623	A1020194	mm5d03 v1 NCI_C6p1	9.55+00
c	959	15	1.3	624	A5172078	mm5d03 v1 NCI_C6p1	9.55+00
c	960	15	1.3	624	B93336	C17-HSP-121886 v1 C17	9.55+00











7	RESULT	N17503	406 bp	ARNA	EST	05-JAN-1998				
	LOCUS	18830	Lambda-PRL2	Arabidopsis thaliana	cDNA clone 206P2277	ARNA				
	DEFINITION	sequence.								
	ACCESSION	N17503								
	VERSION	EST35745								
	KEYWORDS	EST								
	ORGANISM	Arabidopsis thaliana								
	REFERENCE	Eukaryotes: Viridiplantae: Charophyta/Embryophyta group: Embryophyta: Seed Plants: Magnoliopsida: Magnoliopsida: Euphorbiales: Brassicaceae: Arabidopsis 1 (bases 1 to 406)								
	AUTHORS	Newman, T., deBruijn, F. J., Green, P., Kestegat, K., Kende, H., Mintchen, M., Ohlrogge, J., Rakheh, N., Bonerville, S., Thomasow, H., Guns galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)								
	TITLE	95148729								
	JOURNAL									
	COMMENT	Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PRL, Michigan State University Plant Biology Bldg., E. Plant Biology Bldg., 48746-1531 Tel: 517-353-0854 Fax: 517-353-9168 Email: 22313cne@m.cl.msu.edu Seq primer: 77 dya primer. 1 Location/Qualifiers 1 406								
	FEATURES									
	SOURCE									

```

/organs='Arabidopsis thaliana'
/strain='var columbiense'
/clone_lib='Lambada-PRL2'
/site_1='Kip-Low'
/site_2='Not;
quantities of 4 pools of cDNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
of Arabidopsis roots; 3) Arabidopsis root callus; 4)
cycle, half on 16 hr light, 8 hour dark rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
aliquots. This vector is BAC+ lambda ZIP-Lox. The cDNA
library was constructed by ligating the EcoRI-BamHI
019g dt primed cDNA.
/Ab_ref='taxon:3702'
/Ab_ref='Arabidopsis thaliana'
/clone_lib='Lambada-PRL2'
/size=75 c 86 g 137 t 14 others

```

```

small percentage of human cDNAs derived from the human
host cells."
/cd_xref=taxon:5811
/cd_name=gag95ae05.r1"
/cdncs.lib=rgn49 fatchy.oite cdna"
<db><id> All side MRF
<i>>418
BASE COUNT      98 a 139 c   90 g    83 t    8 others
ORIGIN
Quay Match          17%; Score 19; DB 9; Length 418;
Best Local Similarity 95.0%; Pred. No. 250e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
db 368 GCCTATCCTTCGCGGGAC 387
            ||| | | | | | | | | | | | | | | |
oy 1006 GACTTCTTGTCGCGGGAC 1025

```

RESULT	EST
9	
AA014588	435 bp
nm92a01.r1	mouse embryo NM0813.5
clone 440520.5	'similar to PRA:A55093 A55093 fatty acid transport protein precursor - mouse', mRNA sequence.
ACCESSION	94475468
INID	94475468
RETWORDS	EST.
SOURCE	house mouse.
ORGANISM	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE	Basal (bases 1-435).
AUTHORS	Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Stepec, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R.
TITLE	The WABU-HMI Mouse EST Project
JOURNAL	Unpublished (1996)

Contact: Maria Wenzel ESP Project  
WASU-HEM Women's Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 501, St. Louis, MO 63108  
Tel: 314 286 1810  
Fax: 314 286 1810  
Email: mousenzel@wustl.edu  
\*All content is available royalty free through LMR, contact the  
IMAGE resource (lmr@imglib.org) for further information.  
MGI-265556

Query Match	1-78:	Score 19:	DB 12:	Length 406:
Best Local Similarity	90.5%:	Pred.No. 2.50e-06:		
Matches 19:	Conservative:	0:	Mismatches 2:	Indels 0:
Gaps 0:				

db	293	CAGGCTCCTGACATCTCCAG	313			
cp	975	CAGGCTCCTGACATCTCCAG	955			
RESULT	8					
LOCUS	W66185	418 bp	mRNA			EST
DEFINITION	Y828Y91e05.r1 Tgm494 rhychoite cDNA	Toxoplasma gondii cDNA clone				11-SEP-1997
ACCESSION	Y828Y91e05.r1	5'	mRNA sequence.			
KEYWORDS	91374353					
ORGANISM	Toxoplasma gondii.					
SOURCE	Toxoplasma gondii.					
REFERENCE	Sarcocystis gondii; Toxoplasma.					
AUTHORS	1 (bases 1 to 418)					
	Rehli, A., Manger, I., Marra, M., Sidley, L.D., Alioka, J.A.,					
	Wan, K.L., Wotzinger, R. H. and Bochtchay, J.,					
TITLE	WASHU-Merck-Stanford-NIH Toxoplasma EST project					
JOURNAL	Unpublished (1996)					
COMMENT	Contract: Marra N					
	WASHU-Merck EST Project					
	Washington University School of Medicine					
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108					
	Phone: 314 286 1800					
	Fax: 314 286 1810					

```

Email: toxocoston.wustl.edu
Clones are available from Genome Systems (genome@mc.net); the
J. Craig Venter Institute (jcv@venter.com); and the Broad
Institute (broad@broad.mit.edu).
(1)biochempartner.stanford.edu,
Seq primer: T3
High quality sequence stop: 329.
FEATURES
    source
        1..113
            /organism="Toxoplasma gondii"
            /strain="ME49, Clone P25"
            /catalase="yes"
            /site_1="EcoRI; site_2="XhoI;
            /site_3="XhoI;
            /notes="Toxoplasma parasites were grown in human foreskin
            fibroblasts cultures in vitro. The library was
            constructed by A. Behl and I. Menger, Stanford University.
            The library was screened by PCR using the following
            priming and directionally cloned into the EcoRI and XhoI
            sites of the Lambda zap vector using the ZAP-cDNA
            synthesis kit (Stratagene). Warning: the library contains a

```

---

[illegible]

```

M.Felina Bonaldo.*
/db_ref=taxon:10090*
/c clones=40520*
/genus=homo sapiens mouse embryo NMIMZ13.14.5*
/species=homo sapiens
/tissue_type='embryo'
/dev stage='13.5-14.5dpc total fetus'
/lab_host='DBIOB*'
cl.=7435
108 a      106 c    122 g     99 t

Query Match          1.7%: Score 19; DB 23; Length 435;
Best Local Similarity 100.0%; Pred. No. 2.5e+06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

db   jll CAGGAACCCGAGACTTCT 329
        |||||
cp   db   894 CAGGAACCCGAGACTTCT 876

```

RESULT	10	EST	14-MAR-1997
LOCUS	AA254935	457 bp	RNA
DEFINITION	m82h09.1 Soares mouse NKL Mus musculus cDNA clone 719777 5' similar to TR.G563829 G563829 FAT1 ACID TRANSPORT PROTEIN. 1; M82A		
ACCESSION	AA254935		
KEYWORDS	91889521		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	mitochondrial autophagos. Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		







WashU-Merck EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available royalty-free through LHM; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -2843 rev1 5' from Auerham  
High quality sequence stop: 324

# FEATURES

source  
1. 576 bp  
/organism="Homo sapiens"  
/note="Organ: colon; Vector: pBlueScript SK-; site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. T-84 colonic epithelial cell line. Average  
insert size: 1.5 kb. Cloned into XhoI-XbaI adaptor  
vector. 5' adaptor sequence: 5'-GATTCGAGTGGGAG-3';  
3'-adaptor sequence: 5'-GATTCGAGTGGGAG-3';  
/db\_xref="taxon:9606"  
/clone="58486"  
/vector="pBlueScript SK-"  
/lab\_host="SOLiA cells (kanamycin resistant)"  
BASE COUNT 153 a 127 c 125 g 164 t 7 others  
ORIGIN  
1. 78; Score 19; DB 24; Length 576;  
Best Local Similarity 100.0%; Pred. No. 2.50e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 488 CCGAATGGAGGAGTAA 506  
Cp 280 CCGAATGGAGGAGTAA 262

RESULT 15  
LOCUS A1047830 785 bp mRNA EST 08-JUL-1998

DEFINITION uc64b12.x1 Sugano mouse liver m1a Mus musculus cDNA clone 1450655

3' similar to TRP97524 P97524 VERY-LONG-CHAIN ACYL-CoA SYNTHETASE

ACCESSION A1047830

KEYWORDS EST

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 785)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisler, S., Kucan, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Thelander, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0201  
Email: mda@natsig.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tig.org/tig/tgdb/humgen/bac\_end\_search/bac\_end\_search.html  
Class: BAC ends.

# FEATURES

source  
1. 506 bp  
/organism="Homo sapiens"  
/note="Organ: colon; Vector: pBlueScript SK-; site\_1:  
HindIII; Vector: pBlueScript SK-; site\_2:  
/db\_xref="taxon:9606"  
/clone="328815"  
/vector="pBlueScript SK-"  
/sex="Male" CH-ESP  
/cell\_type="Epithelial"  
BASE COUNT 162 a 61 c 98 g 185 t  
ORIGIN  
1. 78; Score 19; DB 27; Length 506;  
Best Local Similarity 100.0%; Pred. No. 2.50e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 CCGAATCACTGAGTAA 57  
Cp 68 CCGAATCACTGAGTAA 50

RESULT 14

LOCUS A1418128 576 bp mRNA EST 05-DEC-1996

DEFINITION uc64b12.x1 Sugano mouse liver m1a Mus musculus cDNA clone

3' similar to TRP97524 P97524 VERY-LONG-CHAIN ACYL-CoA SYNTHETASE

ACCESSION A1418128

KEYWORDS EST

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 576)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisler, S., Kucan, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Thelander, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

QY 876 AGAAGTCTGCTTCCTG 994

RESULT 16

LOCUS A919945 127 bp mRNA EST 20-APR-1998

DEFINITION uc64b12.x1 Sugano mouse liver m1a Mus musculus cDNA clone

3' similar to gb:X54637 NON-RECEPTOR TYROSINE-PROTEIN

ACCESSION X54637

KEYWORDS EST

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 127)

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisler, S., Kucan, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Thelander, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Scheilberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Waterston, R., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.

Waterston, R.



```

BASE COUNT      91 a 49 c 83 g 85 t      1 others
ORIGIN

Query Match      1 6% Score 18: DB 22; Length 311:
Local ID: 1004; Access No: 596-04
Matches: 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 68 TCGACCTGCTGATGTC 85
QY 601 TCGACCTGCTGATGTC 618

RESULT 18 B11064      391 bp      DNA      GSS      15-OCT-1997
DEFINITION      HS-1005-91-D08-MR bp, C17 Human Genomic Screen Library C Homo
                sapiens genomic clone Plate-CT 499 Col=15 Row=H. genomic survey
                sequence.
ACCESSION      X81064
VERSION        B11064.1
KEYWORDS       GSS:
SOURCE         human.
ORGANISM       Homo sapiens
                Chordata; Vertebrata; Mammalia; Eutheria;
                Primates; Catarrhini; Hominoidea; Hominidae; Homo.

REFERENCE      1 (bases 1 to 381)
AUTHORS        Mahairas G.G., Kachrone K.D., Smith T., Tipton S., Schmidt S.,
                Trautman R., Amalin C., Blanchard A., West A. and Hood L.E.
TITLE          Sequencing, Generation and Preliminary Analysis of 20,000 Sequence
                Tagged Connectors
JOURNAL        Unpublished (1997)
COMMENT        Contact: Mahairas GG, Kachrone KD, Hood L
                University of Washington
                Seattle, WA 98195, USA
                Tel: (206) 685-7304
                Fax: (206) 685-7344
                Email: Kachrone@u.washington.edu
                Sequence Tagged Connector
                Plate: CT 499 row: H column: 15
                High quality sequence step: 381.
                Location/Qualifiers
                    1. 381
                        organism="Homo sapiens"
                        db="TDB"
                        E-Cell DB100
                        /db_xref="taxon:9606"
                        /clone="Plate-CT 499 Col=15 Row=H"
FEATURES
    source

```

```
Source: IMAGE Consortium, LMLN  
This clone is available royalty-free through LMLN : contact the  
IMAGE Consortium (info@insdc.llnl.gov) for further information.  
  
FEATURES             location/Qualifiers  
     source           1..403  
                       /organism='Homo sapiens'  
BASE COUNT          119 a   61 c    68 g   155 t  
ORIGIN  
  
Query Match  
Best Local      1.6% Score 18; DB 20; Length 403;  
Query Start     18; Conservat: 0; Mismatch: 0; Indels 0; Gaps 0;  
Db              68 ACCAGGTTTCACATG 85  
|||||            |||||  
Cp              69B ACCAGGTTTCACATG 66I  
  
RESULT 20  
LOCUS       R35098               415 bp        mRNA                ZPT  
DEFINITION  H3K3.3.r1 Homo sapiens CDNA clone 196653.5'.  
ACCESSION   R35098  
KEYWORDS    9737388  
            NID  
SOURCE      Human clone-196653 primer=H3RPI library-Scores fetal liver spleen  
            host=DHIOB (ampicillin resistant) Raitec-Pac I Reito2-Eco RI Liver  
            and spleen from a 20 week-post conception male fetus. 1st strand  
            cDNA was primed with a Pac I - oligopdt primer [5'  
            CCAATTAACTTGAAGAAGAACCGTGCCGGGTGCCTGTCATCATTAAGTACC  
            CGAATAACTTAAAGTAGGTTGGTTTCCCCTTCCCCCTT] , double-stranded  
            cDNA ligated into BamHI-XbaI restriction fragment (Pharmacia).  
            I and cloned into the Pac I and Eco RI sites of the modified pTRV3  
            vector. Library went through one round of normalisation. Library  
            constructed by Bento Soares and M.Patina Bonaldi.  
ORGANISM    Eukaryotes; Metazoa; Euarchonta; Bilateria; Coelomata;  
Euteleostomi; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Actinopterygii; Chonactes; Tetrapoda; Amniota; Mammalia; Theria;  
Placentalia; Archontes; Primates; Catarrhini; Hominoidea; Homo.  
AUTHORS      Hallier,L., Clark,M., Dubouquet,T., Elliston,K., Hawkins,M.,  
Holman,M., Hulsmann,M., Kucaba,T., Le.M., Lennon,C., Marra,M.,  
Parsons,J., Rifkin.L., Rohlfing,T., Soares,M., Tai,P.,  
Wilson,R.B.Z., Watsonson,K., Williamson,A., Wohlmann,P. and  
The WashU-Merck EST Project  
Unpublished 1995)  
CONTACT     Wilton BE  
WashU-Merck EST Project  
JOURNAL     JOURNAL  
COMMENT
```



```

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1800
Email: srs@wustl.edu
High quality sequence stops: 331
Source: IMAGE Consortium, LNCM.
This clone is available royalty-free through LNCM; contact the
IMAGE Consortium (ncicb@lncm.nih.gov) for further information.

FEATURES             1..415
     source            /organism="Homo sapiens"
                        /db="IMAGE" /accession="39653"
                        /map="41..>415"
BASE COUNT          124 a      68 g      161 t      2 others
ORIGIN
>
Query Match          1.69; Score 18; DB 20; Length 415;
Best Local Similarity 100.0%; Pred.No. 1.59e+04;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB    50 ACCAAGCTTTCACATG 67
      |||||
Cp    698 ACCAAGCTTTCACATG 681

RESULT 21
LOCUS      B53590      423 bp      DNA      GSS      19-JUN-1998
DEFINITION C17-HSP-20111.17R C17-HSP Homo sapiens genomic clone 201111,
            genomic survey sequence.
ACCESSION  U55854
VERSION    1
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominoidea; Homo.

REFERENCE  1 (bases 1 to 423)
AUTHORS   Adams M.D., Rounsley S.D., Field C.E., Beas S., Linher K.,
            Stans M.K., Berger J., Kringer D., Suh E., Wible C., Shinya H.,
            et al.
TITLE      Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
JOURNAL    Unpublished (1997)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9700 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208

```

```

Email: mhadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/db/humgen/bac_end_search/bac_end_search.html
Class: BAC ends Reverse
Features:
1. 423
Location/Qualifiers
/organism="Homo sapiens"
/submitter="TIGR"
/submitter_contact="photomk11, Site_1: kindiii, Site_2:
kindiii"
/db_xref="taxon:9606"
/clone_id="70111"
/submitter_contact="TIGR"
/sex="Male"
/cell_type="Sperm"
BASE COUNT 141 a 84 c 75 g 123 t
ORIGIN
Query Match 1.6%; Score 18; DB 27; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.5%-04;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 33 TCAGAAAGTAGCTGCTGA 50
|||||
CP 958 TCAGAAAGTAGCTGCTGA 941
|||||
RESULT 22
LOCUS A1005637 429 bp mRNA 12-JUN-1998
DEFINITION OVS920.01 testes.Testis_MNT Homo sapiens cDNA clone IMAGE:164155
ACCESSION A1005637
NID 93215147
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
MCT (Pages 1 to 429)
AUTHOR MCT, ncbi.nlm.nih.gov/nclogseq.
INSTITUTION National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
CDNA LIBRARY Preparation: M. Bento Soares, Ph.D., M. Fatima
CDNA LIBRARY Arrayed by: Greg Lennon, Ph.D.

```

[illegible]

<b>ORIGIN</b>	Mori I and directionally cloned into the Mori I and Hind III sites of the Lambda HA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.					
<b>ORGANISM</b>	<i>Eumetazoa; Metazoa; Chordata; Vertebrates; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea</i> .					
<b>REFERENCE</b>	1. (bases 1 to 461) Millier, S., Glick, R., Duboucq, T., Elliston, K., Hawkins, M., Williams, J., Johnson, D., Wilson, R., Johnson, A., Barra, H., Trevisan, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, P., and Pavenski, E., Watterston, R., Williamson, A., Wohldmann, P., and Wilson, R., March EST Project The Wellcome Trust Unpublished (1995)					
<b>TITLE</b>						
<b>JOURNAL</b>						
<b>COMMENT</b>	GDB: G100-4897 Contact: Wilson RK Address: Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1800 E-mail: est@wustl.edu High quality sequence stops: 298 Source: IMAGE Consortium, LLNL This clone is available royalty free through LLNL ; contact the LLNL Gene Center ( <a href="mailto:genecenter@llnl.gov">genecenter@llnl.gov</a> ) for further information. Locus Location/Qualifiers 1. .461  organism="Homo sapiens"  cdna="28356"  size=141 g 3 t 5 others					
<b>FEATURES</b>						
<b>SOURCE</b>	BASE COUNT 107 a 115 c 91 t 5 others					
<b>ORIGIN</b>						
<b>QUERY MATCH</b>	Score 18; Score 18; DB 14; Length 461;					
<b>Best Local Similarity</b>	100.0% Identical 100.0%; Mismatches 0; Indels 0; Gaps 0;					
<b>MATCHES</b>	Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
<b>DB</b>	32 OTCCAGACGGCGCTTC 49					
<b>Cp</b>	215 CTCATCACTGCCTTC 198					
<b>PROSITE</b>	PS01556 462 bp mRNA EST					
<b>LOCUS</b>	gdb9607.r1 Soares fetal lung DBHL95 Homo sapiens cDNA clone 308844					
<b>DEFINITION</b>	5' similar to gb:AF1520.406 RBMSOL PROTEIN 624 (RBMN). : mRNA					
<b>SEQUENCE</b>						
<b>ACCESSION</b>	U1101921					
<b>NID</b>	U1101921					



Matches	10	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Db	408	AAAGTGGACGTGAGAT	425						
Cp	522	AAAGTGGACGTGAGAT	505						
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									
29									
30									
31									
32									
33									
34									
35									
36									
37									
38									
39									
40									
41									
42									
43									
44									
45									
46									
47									
48									
49									
50									
51									
52									
53									
54									
55									
56									
57									
58									
59									
60									
61									
62									
63									
64									
65									
66									
67									
68									
69									
70									
71									
72									
73									
74									
75									

Contact: Wilson KE  
WUHL Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel.: 314 286 1800  
Fax: 314 286 1897  
Email: ke@seismo.wustl.edu

Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([llnl.llnwd.gov](mailto:llnl.llnwd.gov)) for further information.

FEATURES source  
1..477  
/organism=Homo sapiens\*  
/clone=J1473  
BASE COUNT 109 A 85 C 115 G 155 T 12 others  
ORIGIN

Query Match 1.48; Score 18; DB 14; Length 477;  
Local similarity 97%; Position 89-144;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DD 318 CTGTCGGCTGTCTATTTT 336  
          |||||  
Cc 551 CTGTCGGCTGTCTATTTT 533

RESULT 27 AS131233 509 bp mRNA EST 20-AUG-1997  
n152e06.al NCI\_CGAP\_P9 Homo sapiens cDNA clone IMAGE:596130, MNNA  
sequence.  
ACCESSION AS131233  
KEYWORDS EST, P339  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Mammalia; Hominidae, Primates, Catarrhini, Hominoide;  
Homo.

REFERENCE 1 (bases 1 to 509)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
Human Genome Project, Institute, Cancer Genome Anatomy Project (CGAP).  
TITLE Human Genome  
COMMENT Published (1997)

Contact: Robert Strausberg, Ph.D.  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo F. Chuquill,  
M.D., Michael R. Emmert-Buck, M.D., Ph.D.







Iteration: R  
 TITLE  
 The MASHU-HMI Mouse EST Project  
 JOURNAL  
 Unpublished (1996)  
 COMMENT  
 Contact: Maura M/Mouse EST Project  
 mma80@hmi.house.esci.net  
 MASHU-HMI Mouse EST Project  
 Department of Molecular and Medical  
 Genetics  
 4444 Forest Park Parkway, Box 5501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseactivation.wustl.edu  
 This clone is available royalty-free through MML; contact the  
 MML: 20895atlum (info@mls.lim.gov) for further information.  
 MGI: 20895atlum  
 Trace considered overall poor quality  
 Seq primer: ETP-primers  
 High quality sequence stop: 1.  
 Location/Qualifiers

```
/organism='Mus musculus'
/note=Vector: pF773D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; left strand cDNA
insertion site: NotI-EcoRI linker oligo(CT) primer [5'-GTTTGGTTC-3'];
TGATCAATTACAGCGCCGCAGCAGCCCCGCCAAGAAGGAAAGTAAGGTGAATA
double-stranded cDNA was also selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pF773 vector
(Pharmacia). Library went through one round off by Benito
Soares and M.Patrina Bonaldi. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
/_AB_xref=taxon:10090
/_clone=_library:_EcoRI_Soares mouse plasmid 9.5"
/_comment=_Library:_EcoRI_Soares mouse plasmid 9.5"
/_date_stage=19.5 dpc total fetus"/
/lab_host=DBLOB (ampicillin resistant)"
      55   a    74   c     65 t
          1 5% ; Score 17; DB 10; Length 253;
Query Match Best Local Similarity 100.0%; Pred.No. 8.05e-03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Tue Nov 17 08:55:30 1998

```

/db_xref=taxon:7955"
/cdname="l38"
/clone_lib="zebrafish ICRFsls"
/issue_type="pooled 26-somite embryos"
/lab_host="XL-blue MFP"
cl.->262
76 a 38 c 57 g 91 t

1.5b: Prod 17: DB 7: Length 262:
Meth: Local Similarity 100.0%: Sred 16: 8.05e-01:
Machine 17: conservative 0: Hammetts 0: Indels 0: Gaps 0:

db 180 TTCACGAAATACCA 206
py 526 TTCACGAAATACCA 542

RESULT 33 AA492112 264 bp sRNA 15-NOV-1987
DEFINITION seq1401.a1 NC_049412 Homo sapiens cDNA clone IMAGE:298655, sRNA
ACCESSION AA492112
VERSION 1.0
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Notes: Mitochondrial subunit 6, Metazoa, Chordata,
Vertebrates, Mammalia, Euthera, Primates, Catarrhini, Hominoideae;
Homo.

```

**JOURNAL**  
**TITLE** NC-CGAP NIDH//www.ncic.nih.gov/ncicgap.  
**AUTHORS** NC-CGAP NIDH//www.ncic.nih.gov/ncicgap.  
**ABSTRACT** Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph. D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emert-Buck, M.D., Ph. D.  
CDNA Library: David B. Krisman, Ph. D.  
DNA sequencing by Washington University Genome Sequencing Center  
clone distribution: NC-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LUNH at:  
www.bio.lnlni.gov/dbfp/lsange/lsange.html  
Insert Length: 1086 Std Error: 0.00  
Seq primer: -40ml fwd. RT from Amersham



```

mice identifies a large number of male germ cell-specific sequence
JOURNAL      J6201. Reprod. 52 (1), 131-138 (1995)
MEDLINE      95226550
REFERENCE    2 (bases 1 to 287)
AUTHORS     Direct Submission
TITLE       Submitted (1-MAR-1994) C.Hoog, Laboratory of Molecular Genetics,
            Karolinska Institute, 171 77 Stockholm, Sweden
FEATURES    /Accession/Qualifiers
            /organism="Mus musculus"
            /strain="CEB"
            /db_xref="taxid:10090"
            /date_rec="19940311"
            /tissue_type="Testis"
            /clone_lib="T-SAP"
BASE COUNT   67 a  81 c  66 g          73 t
ORIGIN
Query Match           1:58; Score 17; DB J3; Length 287;
Best Local Similarity 100.0%; Pred. No. 8.0Se-03;
Matches              0; Conservative 0; Mismatches 0; Gaps 0; Caps 0;

Db 131 CNTVCCATTCGCGCA 147
Qy 266 CNTCCTCATTCGCGCA 282
|||||
|||||

RESULT 36
LOCUS       T02045             289 bp mRNA
DEFINITION T02045 Mus musculus Cerebellar Agenesis cDNA clone CE25163 similar to
            Tropoin I.
ACCESSION   T02045
VERSION     1
KEYWORDS    T02045
SOURCE      EST.
            Genbank clones=CE25163 library=Early embryo, Stratagene (cat.
            #937/007) primer=M31 Reverse.
ORGANISM   Eucariotae; Animalia; Eumetazoa; Nematoda; Secernentea; Rhabditida;
            Nematode; Caenorhabditis elegans
REFERENCE   1 (bases 1 to 289)
            McCombs,W.R., Kelley,J.M., Aubin,L., Goscochea,M.,
            Fitzcerald,M.G., Wu,A., Adams,M.D., Dubnick,M., Keriavage,A.H.,
            Venter,J.C. and Fields,C.A.
            Unpublished (1993).
JOURNAL
COMMENT     Contact: Keriavage AH
            The Institute for Genomic Research
            9700 Medical Center Drive, Gaithersburg, MD 20878
            Tel.: 301/662-9058

```

```

source
1, 296
/notes="Homo sapiens"
/organ="Organ: Liver and Spleen; Vector: p7773D (Pharmacia)
1st strand cDNA was primed with a Pac I - oligo(dT) primer
(Pharmacia), digested with EcoRI and ligated to EcoRI-doped,
double-stranded cDNA was ligated to EcoRI-doped cDNA
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7773 vector. Library
went through one round of normalisation. Library
deposited at EMBL, accession number AF081242."
/Ab_xref=taxon:9606
/clonelib="Scares fetal liver spleen mRFLS"
/dev_stage="20 week-post conception fetus"
/lab_host="DHIO3 (ampicillin resistant)"
89 a 67 c 71 g 68 t 1 others
Query Match 1.58; Score 17; DB 22; Length 296;
Best Local Similarity 100.0%; Pred. No. 8.05e-03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 27 TCTGTCGTTCGGCCGC 43
|||||
QY 140 TCTGTCGTTCGGCCGC 156
RESULT 38
A1017132 301 bp mRNA EST 16-JUN-1998
DEFINITION Homo sapiens CDNA clone AB02:103191
3' mRNA sequence.
ACCESSION A1017132
MIM 9321468
NID 9321468
SOURCE human
ORGANISM
Homo sapiens
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
KCI-CGAP http://www.ncbi.nlm.nih.gov/kci/cgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
```











Source

/organism=Homo sapiens"  
NucleoAcidSeq: p773Dpc (Pharmacia), with a modified  
restriction site at the 3' end (GGGAGCG) and cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20<sup>+</sup>, IgD<sup>-</sup>,  
IgM<sup>+</sup>) [Saito et al., 1986]. The PCR product was purified  
using a QIAquick PCR Purification kit (Qiagen Inc.) and  
Mr. Gerald Hart (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer:  
1'. TGGACATCTCAAGCCGACCCTGCCTCATTCCTTTTCCCC-  
3'. The PCR product was digested with Not I and cloned into the Not I  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p773 vector. Library  
construction and screening were performed as described  
constructed by Bento Soares and M. Fatima Bonaldi.

```

/Ab_xref.*taxon:9606*
/clones.*IMAGE:712497*
/clone.lib.*NCI_CGAP_GCB1*
/tissue_type.*terminal center B cell*
/lab_host.*DH08*
complement(<1., >141)
104 a 48 c 42 g 147 t
mRNA
BASE COUNT
ORIGIN

```

```

Query Match      Score 17: DB 8: Length 341:
Best Local Similarity 100.0%:
Pred. No. 8.05e-03:
Matches 17: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
bb 275 ACTTGTCTCAACCA 291
      |||||

```

46	CGS	03-AUG-1988
RESULTS	DNA	
ACCESION	345 bp	
DESCRIPTION	NP 2245, 2.605 kb	
	septans genomic clone Plate-2245 Col-18 Row-M, genomic survey	
	sequence.	
ACCESION	0066765	
DATE	01-07-83	
KEYWORDS	SSS	

Page 76

Sarcotrypani; Chonotata; Tetrapoda; Amniota; Mammalia; Theria;  
 Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
 (Cases 1 to 352)  
 Holman, A., Putnam, M., Kuchel, T., & M. Lennon, G. M., et al.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, H., Tan, F., et al.,  
 Trevisan, E., Watzerson, R., Williamson, A., Wohlmann, P., and  
 Wilson, R.  
 The Human Genome Project  
 (Unpublished 1995)

Contact: Wilson R. Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1810  
Fax: 314 286 1810  
Email: wproject@wustl.edu  
Source: IMAGE Consortium  
This clone is available royalty-free through IMAGE. Contact the  
IMAGE Consortium for more information.

FEATURES	source	BASE COUNT	ORIGIN
IMAGE CONJUNCTION (info@image.iuh.gov) for further information.			
Location/Qualifiers			
1. .352			
/organism="Homo sapiens"			
/clone="15025"			
55 a	97 c	104 g	92 t
			4 others

```

Query Match      i.5%  Score 17: DB 20: Length 352:
Query Local Similarity 100.0%:  Pred. No. 8 0.0e+00:
               0: Mismatches 0: Indels 0: Gaps 0:
               17: Conservative
bb  60  TCTCTGACGCTGCT 76
      | | | | | | | | | |
yy  889  TCTCTGACGCTGCT 905

```

Sarcotrypani; Chonotata; Tetrapoda; Amniota; Mammalia; Theria;  
 Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
 (Cases 1 to 352)  
 Holman, A., Putnam, M., Kuchel, T., & M. Lennon, G. M., et al.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, H., Tan, F., et al.,  
 Trevisan, E., Watzerson, R., Williamson, A., Wohlmann, P., and  
 Wilson, R.  
 The Human Genome Project  
 (Unpublished 1995)

Contact: Wilson R. Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1810  
Fax: 314 286 1810  
Email: wproject@wustl.edu  
Source: IMAGE Consortium  
This clone is available royalty-free through IMAGE. Contact the  
IMAGE Consortium for more information.

FEATURES	source	BASE COUNT	ORIGIN
IMAGE CONJUNCTION (info@image.iuh.gov) for further information.			
Location/Qualifiers			
1. .352			
/organism="Homo sapiens"			
/clone="15025"			
55 a	97 c	104 g	92 t
			4 others

```

Query Match      i.5%  Score 17: DB 20: Length 352:
Query Local Similarity 100.0%:  Pred. No. 8 0.0e+00:
               0: Mismatches 0: Indels 0: Gaps 0:
               17: Conservative
bb  60  TCTCTGACGCTGCT 76
      | | | | | | | | | |
yy  889  TCTCTGACGCTGCT 905

```



LOCUS	AA260146	353 bp	BBNA	EST	18-APR-1987
DEFINITION	v37605.r1 Soares mouse JNK212 5' Mus musculus cDNA clone 733568 5'				
ACCESSION	M584 sequence.				
VERSION	AA260146				
KEYWORDS	EST 736031				
ORGANISM	house mouse.				
REFERENCE	Mus musculus				
AUTHORS	1 (bases 1 to 353)				
	Carroll, H. L., Allen, M., Bowles, M., Dietrich, N., Dubouquet, T., Schellenberg, R., Steptoe, M., Tan, P., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R. Mouse EST Project				
TITLE	JNK212 mouse EST Project				
COMMENT	Duplicated (1996)				
	Contact: Maria M/ Mouse EST Project				
	Washington University Project				
	4444 Forest Park Parkway, Box 5501, St. Louis, MO 63108				
	Tel: 314 286 1800				
	Fax: 314 286 1810				
	This clone is available royalty-free through MAMU: contact the IMAGE Consortium (infoimage@iml.gov) for further information.				
	MGI:450616 - 28kb rev20 EST from Amersham				
	High quality sequencing: 100.				
FEATURES	Location/Qualifiers				
source	1..353				
	/organism="Mus musculus"				
	/vector="pTZ19-pac (Pharmacia) with a modified				
	polylinker. Site 1: Not I. Site 2: Eco RI. 1st strand				
	was primed with a Not I - oligo(dT) primer [5']				
	on total mouse RNA (provided by Mironov to Wayne State				
	university) [5']				
	Adaptors (Pharmacia), digested with Not I and cloned into				
	the Not I and Eco RI sites of the modified pTZ19 vector.				
	The cDNA was sequenced using the primer provided and was				
	constructed by Bento Soares and M. Patricia Bonaldo.				
	/db_xref="taxon:10090"				
	/clone="733568"				
	/sex="unknown"				
	/dev_stage="12 dpoc total fetus"				

FEATURES	108 a	60 c	73 g	109 t	3 others		
source	/organism="Homo sapiens" /clone="149539"						
BASE COUNT							
ORIGIN							
Query Match	1.5%: Score 17; DB 14; Length 353;						
Best Local Similarity	100.0%; Pred. No. 8.0e-03;						
Matches	17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
DB	299	ACACCTCTTCTTCTT	315				
Qy	248	ACACCTCTTCTTCTT	264				
RESULT 50							
LOCUS	360 bp mRNA						
DEFINITION	C50268 has cDNA clone yk47925 : 5' end, single read, mRNA sequence.						
ACCESSION	U02387.519						
KEYWORDS	EST; EST(expressed sequence tag);						
REMARKS	Cenorhabditis elegans (strain N2) embryo hemaphroditic embryo cDNA						
SOURCE	clone yk47925						
ORGANISM	Cenorhabditis elegans						
REFERENCE	Eukaryotes; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;						
AUTHORS	1 (singles); Rhabditoidae; Rhabditidae; Polidoriinae; Cenorhabditis.						
TITLE	Kohara, Y., Morishashi, T., Tabara, H., Shin, I.-T., Watanabe, H.,						
REFERENCE	Sano, M., Miyata, A., Ohta, T., Mitani, T., Tseng, H., Sugura, I.,						
AUTHORS	Serafini, Sugimoto, K., Iida, K. and Washigaki, A.						
JOURNAL	Unpublished (1997)						
REFERENCE	the C. elegans genome						
AUTHORS	2 (bases 1 to 360)						
JOURNAL	Kohara, Y., Kamigaito						
REFERENCE	Submitted (23-JUN-1997) to the DDBJ/EMBL/GenBank databases, Yuiji						
AUTHORS	Kohara, National Institute of Genetics, Gene Network Lab; Yata						
JOURNAL	111, Mishima, Shizuoka 411, Japan (E-mail: yk@haraelab.nig.ac.jp,						
REFERENCE	Tel.: 01-359-8520; Fax: 01-359-81-6855						
AUTHORS	/organism="Homo sapiens"						
JOURNAL	/clone="149539"						
REFERENCE	1. 360						
AUTHORS	/organism="Cenorhabditis elegans"						
JOURNAL	/strain="N2"						
REFERENCE	/accession="U02387.519"						
AUTHORS	/clone="yk47925"						
JOURNAL							
REFERENCE							
AUTHORS							
JOURNAL							
REFERENCE							
AUTHORS							
JOURNAL							
REFERENCE							
AUTHORS							
JOURNAL							
REFERENCE							
AUTHORS							
JOURNAL							
REFERENCE							
AUTHORS							
JOURNAL							
REFERENCE							
AUTHORS							
JOURNAL							
REFERENCE							
AUTHORS							
JOURNAL							
REFERENCE							
AUTHORS							

Tue Nov 17 08:55:30 1998

```

      mRNA      /lab host="DH10B"
BASE COUNT   100 a      90 g      82 t
ORIGIN
<1> 1-253

Query Match      1.5%:      Score 17: DB 26; Length 333;
Best Local Similarity 100.0%:      Pred. No. 0.95e-03; Caps 0;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Db 120 TGAAGCATCTTGAGGAC 136
QT 956 TGAAGCATCTTGAGGAC 972

RESULT 49
LOCUS      H00187      353 bp      mRNA      EST      19-JUN-1995
DEFINITION Y12320.r1 Homo sapiens cDNA clone 149539 5'.
ACCESSION  F0643220
KEYWORDS   EST.
SOURCE     human cDNA=149539 library=Soares placenta N02HP vector=pfT71D
           (Pharmacia) with a modified polynucleotide DH10B (ampicillin
           resistance) host=K12. The cDNA was prepared from placenta
           obtained at birth (full term). 1st strand cDNA was primed with a
           Not I - oligo (GT) primer [5',
           AACTGAGGAAATTCGCGCCGACGAGATTTTTTTTTTTTTTTT 3'], double-stranded
           cDNA was ligated to 29p X adaptors (Pharmacia), digested with Not
           I, ligated to the 29p X adaptors, and ligated to the N02HP vector.
           The library went through one round of normalization. Library
           constructed by Bento Soares and M.Patima Bonaldi.
           Homo sapiens Mesencephalon; Eumetazoa; Bilateria; Coelomata;
           Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
           Sarcopterygii; Choroanata; Tetrapoda; Amniota; Mammalia; Theria;
           Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
           Hillier, L., Clark, M., Dubucque, T., Eilenson, K., Hawkins, M.,
           Holman, M., Rulman, M., Dubucque, T., Le, M., Lennon, C., Marra, M.,
           Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
           Wierzbicki, E., Westcotton, R., Williamson, A., Wohlman, P., and
           The WashU-Werch EST Project
           Unpublished (1995)

ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT
Contact: Wilson RK
WashU-Werch EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1310
Email: estest@wustl.edu

```

Tue Nov 17 08:55:30 1998

```

/home/libe/Tuji_Kobura/unpublished/cDNA_library*
/dev/stage=embryo
/sec=hermaprodite*
/tissue_type=embryo

BASE COUNT      94 a      85 c      82 g      85 t
ORIGIN
Query Match      1.54; Score 17; DB 9; Length 360;
Similarity 100; Ids 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0; Indels 0;
Db      54 ACTGCTTCCGAGAAA 70
          TTTTCTTTT
Cp      1056 ACTGCTTCCGAGAAA 1040

Search completed: Sat Nov 14 03:05:54 1998
Job time : 3240 secs

```







```

CC the effects of chemotherapy.
QQ Sequence 360 AA:
Query Match
  Best Local Similarity 100.0%: Pred. No. 2.09e-03; Length 360;
  Hits 71 conservative 0; Mismatches 0; Indels 0; Caps 0;
Db 71 rantsdylln 80
  | | | | | | | |
Qy 69 RNSMDYLLN 78
RESULT 3
ID W02688 standard; peptide; 312 AA.
CD 12698
LN 13-80V-1996 (first entry)
DE G-protein coupled human interleukin-8 receptor.
KW G-protein coupled receptor; ligand binding assay; transmembrane domain;
RN 12698
RT 12698
TX 12698
ZZ 12698
ID W02688 standard; peptide; 312 AA.
CD 12698
LN 13-80V-1996 (first entry)
DE G-protein coupled human interleukin-8 receptor.
KW G-protein coupled receptor; ligand binding assay; transmembrane domain;
RN 12698
RT 12698
TX 12698
ZZ 12698

```

GR HES-60 antihepatoma.  
PR 18-MAR-1996  
PP 10-SEP-1992: 943236.  
PP 10-SEP-1992: US-943236.  
PR 09-SEP-1993: US-118270.  
PR 09-SEP-1993: US-118270.  
PR MURPHY RV SCHWARTZ STAFFE.  
WI: 96-04785/721.  
WP1: 96-10845 receptor peptide - useful as antipsychotic agent, e.g. for treating schizophrenia  
PR 18-APR-1996: US-943236; 108;  
CC G-protein coupled receptor  
CC (GPR) proteins selected from cAMP adenosine nucleotinic acetylcholine,  
CC adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin,  
CC vorarant, cytomagaloviral and other GPR proteins. The receptor proteins  
CC were used to design polypeptides, pref. based on the transmembrane  
CC domain of the protein coupled receptor binding assays.  
CC The polypeptide fragments are used for receptor binding assays.  
CC The polypeptide fragments are used for receptor binding assays.  
CC GPR ligand or modulating GPR ligand binding to a GPR (see WO274-W0299  
CC for examples of polypeptide fragments). The polypeptide fragments can  
CC be used in compositions for treating subjects suffering from a pathology  
CC related to a GPR abnormality e.g. a psychotic disorder such as  
CC schizophrenia. 312 MA;  
CC Sequence

Query Match 2.5% Score 9; DB 19; Length 312;  
Best Local Similarity 100.0%; Pred. No. 4.48e-02;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

26-JUL-1993 (first entry)  
Human IL-8 receptor from clone p2.  
Interleukin-8 receptor; probes; gene therapy; gpc receptor;  
intracellular-calcium mobilising; ligand-binding; MID-2 receptor.  
Homo sapiens.

[illegible]

RESULT 6  
ID R65188 standard; Protein; 356 AA.  
AC R65188:

DT 19-APR-1995 (first entry)  
DE Murine mu-subtype opioid receptor.  
KW Mu-subtype opioid receptor; MSOR; drug addiction.  
OS Rattus rattus

key	modified_site	Location/Qualifiers
PT	10..12	/note= "Putative N-linked glycosylation site"
PT	230	/note= "Threonine residue especially favourable"







PI Chunchazapai A., Hebert C., Kim KJ, Lee J;  
DR WPI: 95-28151/37.  
DR R-PSDB: 095008.  
CC Recombinant mammalian interleukin-8 receptor - used to treat  
or prevent inflammation, also for detecting receptor expression and  
PI purification.  
CC Dictionaire: Columns 53-56; 62pp; English.  
CC Antibodies directed against the interleukin-8 receptor B can be used  
CC to detect inflammation, also for detecting receptor expression and  
CC rheumatoid arthritis and particularly inflammatory bowel disease and  
CC chronic lung inflammation. When immobilised, these antibodies may  
CC be used to detect interleukin-8 receptor B expression in cells and  
CC tissues and for affinity purification of interleukin-8 receptor B  
CC from cells. 360 AA;  
SQ Sequence 360 AA;

Query Match 2.5%; Score 9; DB 15; Length 360;  
Best Local Similarity 100.0%; Pred. No. 4.48e-02;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 314 Yafiqgkfr 322  
|||||

QY 307 YAFIQGKFR 315  
|||||

RESULT 12 standard; Protein: 398 AA.  
AC R1964: 1995 (first entry)  
DE 20-OCT-1995 (first entry)  
DE Rat mu opiate receptor.  
KW Mu opiate receptor; MOR-1; gene therapy, diagnostic.  
RN W0507981-A.  
PN W0507981-A.  
PD 23-MAR-1995.  
PF 13-SEP-1994; U10358.  
PR 13-SEP-1994; U10358.  
PI Tu L; ) DRIV INDIANA FOUND.  
DR R-PSDB: 089222.  
DR WPI: 95-131351/17.  
CC Recombinant acid encoding new human mu opiate receptor - and  
CC related vectors, transformed cells, antibodies etc., useful in  
CC diagnosis, treatment and drug screening.  
CC Dictionaire: Page 190-194; 265pp; English.  
CC A 365 bp fragment of the mouse delta opiate receptor was used to  
CC screen a rat cDNA library under conditions of high stringency.  
CC One positive clone included the coding sequence of the delta  
CC mu opiate receptor, MOR-1 (R71964). MOR-1 was stably expressed in  
CC transfected CHO cells.  
SQ Sequence 398 AA;

PI Recombinant mammalian interleukin-8 receptor - used for screening  
PI interleukin-8 binding antagonists, used to treat inflammation  
PI (R1964: 1995 (first entry).  
CC Peritoneal neutrophils and used as a source of poly(A)+ RNA, to  
CC produce a rabbit neutrophil cDNA library. 250,000 recombinant  
CC plaques were screened for those which hybridized to an antisense  
CC sequence derived from the second transmembrane domain of the  
CC coupled receptors. After tertiary screening, six plaques were  
CC isolated. The insert of one of these plaques, termed F3R was of 2.5  
CC kb in size. This insert was sequenced. The protein deduced from  
CC the sequence showed homology to the second transmembrane domain of  
CC G-protein-coupled receptors. The deduced protein sequence  
CC indicates seven putative transmembrane segments. A human  
CC peripheral blood leukocyte lambda gtl cDNA library (5' stretch) was  
CC screened with the 652 bp EcoRI/HindIII fragment (including nucleotides  
CC 275-925) of F3R. Several human clones which hybridized to the rabbit IL-8 probe  
CC were isolated. The insert of one such clone, termed 4A8 was  
CC sequenced (Q30012).  
SQ Sequence 380 AA;

Query Match 2.5%; Score 9; DB 6; Length 360;  
Best Local Similarity 100.0%; Pred. No. 4.48e-02;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 314 Yafiqgkfr 322  
|||||

QY 307 YAFIQGKFR 315  
|||||

RESULT 11 standard; Protein: 360 AA.  
AC R80758: 1995 (first entry)  
DE 20-OCT-1995 (first entry)  
DE Interleukin 8 receptor B  
KW Rheumatoid arthritis; inflammatory bowel disease;  
KW chronic lung inflammation; treatment; antibody;  
OS Homo sapiens.  
PN U55440011-A.  
PD 08-AUG-1995. 677211.  
PR 28-MAR-1994; U727211.  
PR 28-FEB-1994; U5202036.  
PA (CHUN)/ CHUNTHARAPAI A.  
PA (HEBE)/ HEBERT C.  
PA (KIM)/ KIM K J.  
PA (LEE)/ LEE J.

PD 23-MAR-1995.  
PR 13-SEP-1994; U10358.  
PR 13-SEP-1993; U5-120601.  
PA (INDV ) UNIV INDIANA FOUND.  
DR WPI: 95-131351/17.  
DR R-PSDB: 089222.  
PT New nucleic acid encoding new human mu opiate receptor - and  
PT related vectors, transformed cells, antibodies etc., useful in  
PT diagnosis, treatment and drug screening.  
PS Claim 4: Page 211-214; 265pp; English.  
CC A cDNA library constructed from human caudate nucleus mRNA was  
CC screened with rat mu opiate receptor cDNA under conditions of  
CC high stringency. One positive clone included the sequence given in  
CC Q89226. This clone was sequenced and found to encode the delta  
CC mu opiate receptor, MOR-1 (R71964). MOR-1 was stably expressed in  
CC transfected CHO cells.  
SQ Sequence 400 AA;

Query Match 2.5%; Score 9; DB 13; Length 400;  
Best Local Similarity 100.0%; Pred. No. 4.48e-02;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 332 clnplvlyaf 340  
|||||

QY 301 CLNPLVLYAF 309  
|||||

RESULT 15 standard; Protein: 400 AA.  
AC R76780: 1995 (first entry)  
DE 11-DEC-1995 (first entry)  
DE Human mu opiate receptor.  
KW Mu opiate receptor; MOR; opiate agonist; opiate antagonist;  
RN W0507981-A.  
PN W0507981-A.  
PD 23-MAR-1995.  
PF 13-SEP-1994; U10358.  
PR 13-SEP-1993; U5-120601.  
PA (INDV ) UNIV INDIANA FOUND.  
DR WPI: 95-131351/17.  
DR R-PSDB: 089222.  
PT New nucleic acid encoding new human mu opiate receptor - and  
PT related vectors, transformed cells, antibodies etc., useful in  
PT diagnosis, treatment and drug screening.  
PS Claim 4: Page 211-214; 265pp; English.  
CC A cDNA library constructed from human caudate nucleus mRNA was  
CC screened with rat mu opiate receptor cDNA under conditions of  
CC high stringency. One positive clone included the sequence given in  
CC Q89226. This clone was sequenced and found to encode the delta  
CC mu opiate receptor, MOR-1 (R71964). MOR-1 was stably expressed in  
CC transfected CHO cells.  
SQ Sequence 400 AA;

Query Match 2.5%; Score 9; DB 13; Length 400;  
Best Local Similarity 100.0%; Pred. No. 4.48e-02;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 332 clnplvlyaf 340  
|||||

QY 301 CLNPLVLYAF 309  
|||||

RESULT 15 standard; Protein: 400 AA.  
AC R76780: 1995 (first entry)  
DE 11-DEC-1995 (first entry)  
DE Human mu opiate receptor.  
KW Mu opiate receptor; MOR; opiate agonist; opiate antagonist;  
RN W0507981-A.  
PN W0507981-A.  
PD 23-MAR-1995.  
PF 13-SEP-1994; U10358.  
PR 13-SEP-1993; U5-120601.  
PA (INDV ) UNIV INDIANA FOUND.  
DR WPI: 95-131351/17.  
DR R-PSDB: 089222.  
PT New nucleic acid encoding new human mu opiate receptor - and  
PT related vectors, transformed cells, antibodies etc., useful in  
PT diagnosis, treatment and drug screening.  
PS Claim 4: Page 211-214; 265pp; English.  
CC A cDNA library constructed from human caudate nucleus mRNA was  
CC screened with rat mu opiate receptor cDNA under conditions of  
CC high stringency. One positive clone included the sequence given in  
CC Q89226. This clone was sequenced and found to encode the delta  
CC mu opiate receptor, MOR-1 (R71964). MOR-1 was stably expressed in  
CC transfected CHO cells.  
SQ Sequence 400 AA;

Query Match 2.5%; Score 9; DB 13; Length 398;  
Best Local Similarity 100.0%; Pred. No. 4.48e-02;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 330 clnplvlyaf 338  
|||||

QY 301 CLNPLVLYAF 309  
|||||

RESULT 13 standard; Protein: 398 AA.  
AC R76781: 1995 (first entry)  
DE 11-DEC-1995 (first entry)  
DE Mu opiate receptor; MOR; opiate agonist; opiate antagonist;  
KW drug abuse; analgesic.  
OS Homo sapiens.  
PN U55440011-A.  
PD 08-AUG-1995. 677211.  
PR 28-MAR-1994; U727211.  
PR 28-FEB-1994; U5202036.  
PA (CHUN)/ CHUNTHARAPAI A.  
PA (HEBE)/ HEBERT C.  
PA (KIM)/ KIM K J.  
PA (LEE)/ LEE J.

Query Match 2.5%; Score 9; DB 14; Length 398;  
Best Local Similarity 100.0%; Pred. No. 4.48e-02;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 330 clnplvlyaf 338  
|||||

QY 301 CLNPLVLYAF 309  
|||||

RESULT 14 standard; Protein: 400 AA.  
AC R71966: 1995 (first entry)  
DE 20-OCT-1995 (first entry)  
DE Mu opiate receptor; MOR; gene therapy; diagnostic.  
OS Homo sapiens.  
PN W0507981-A.



```

Query Match      2.5%  Score 9; DB 14; Length 400;
Best-Local Similarity 100.0%;  Pred. No. 4,486-02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 332 clnp1yaf 340
      |||||
QY 301 GNP212IN 309

```



.....  
MPSQLH  
.....  
(TM)  
.....

Release 3.1A John F. Collins, Bioinformatics Research Unit,  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MParch\_Pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Nov 16 10:37:45 1998; MacPar time 14.65 Seconds  
910.311 Million cell updates/sec

Tabular output not generated.

Title: >US-08-887-977-10  
Description: (1-365) from US08887977.ppf  
Perfect Score: 365  
Sequence: 1 NPSVPAIIQSLHNTQL.....NISRQSTADNDNRASFTM 365  
Scoring table: TABLE uniprotatable  
GAP 60

Searched: 120441 seqs, 36531193 residues  
Post-processing: Minimum Match 03  
Listing first 100 summaries

Database: p1r36  
1.p1r2: p1r2 3.p1r3 4.p1r4 5.p1r5

Statistics: Mean 3.585; Variance 0.459; scale 7.805

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Match	Length	ID	Description	Pred. No.
1	315	86.3	369	2	US-08-887-977-10.rpx	0.00e+00

53	7	1.9	648	2	A54892	Na <sup>+</sup> -dependent nucleoside	1.19e+00
54	7	1.9	690	2 <th>A54211</th> <th>ATM1 protein precursor</th> <th>1.19e+00</th>	A54211	ATM1 protein precursor	1.19e+00
55	7	1.9	704	2 <th>B45506</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	B45506	H <sup>+</sup> -transporting ATPase	1.19e+00
56	7	1.9	719	2 <th>B45755</th> <th>hypothetical protein</th> <th>1.19e+00</th>	B45755	hypothetical protein	1.19e+00
57	7	1.9	734	2 <th>B45755</th> <th>hypothetical protein</th> <th>1.19e+00</th>	B45755	hypothetical protein	1.19e+00
58	7	1.9	948	1 <th>PM002</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	PM002	H <sup>+</sup> -transporting ATPase	1.19e+00
59	7	1.9	948	2 <th>S52739</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	S52739	H <sup>+</sup> -transporting ATPase	1.19e+00
60	7	1.9	949	1 <th>PM003</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	PM003	H <sup>+</sup> -transporting ATPase	1.19e+00
61	7	1.9	949	2 <th>PM003</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	PM003	H <sup>+</sup> -transporting ATPase	1.19e+00
62	7	1.9	949	3 <th>PM003</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	PM003	H <sup>+</sup> -transporting ATPase	1.19e+00
63	7	1.9	952	1 <th>S50752</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	S50752	H <sup>+</sup> -transporting ATPase	1.19e+00
64	7	1.9	952	2 <th>S33548</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	S33548	H <sup>+</sup> -transporting ATPase	1.19e+00
65	7	1.9	956	2 <th>A43637</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	A43637	H <sup>+</sup> -transporting ATPase	1.19e+00
66	7	1.9	956	3 <th>S50752</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	S50752	H <sup>+</sup> -transporting ATPase	1.19e+00
67	7	1.9	956	4 <th>S50752</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	S50752	H <sup>+</sup> -transporting ATPase	1.19e+00
68	7	1.9	957 <td>1<th>A41779</th><th>H<sup>+</sup>-transporting ATPase</th><th>1.19e+00</th></td>	1 <th>A41779</th> <th>H<sup>+</sup>-transporting ATPase</th> <th>1.19e+00</th>	A41779	H <sup>+</sup> -transporting ATPase	1.19e+00
69	7	1.9	1028 <td>2<th>G44595</th><th>acetylcholine receptor</th><th>1.19e+00</th></td>	2 <th>G44595</th> <th>acetylcholine receptor</th> <th>1.19e+00</th>	G44595	acetylcholine receptor	1.19e+00
70	7	1.9	1402 <td>2<th>S62557</th><th>hypothetical protein</th><th>1.19e+00</th></td>	2 <th>S62557</th> <th>hypothetical protein</th> <th>1.19e+00</th>	S62557	hypothetical protein	1.19e+00
71	7	1.9	1402 <td>3<th>S62557</th><th>hypothetical protein</th><th>1.19e+00</th></td>	3 <th>S62557</th> <th>hypothetical protein</th> <th>1.19e+00</th>	S62557	hypothetical protein	1.19e+00
72	6	1.6	81	2 <th>S28741</th> <th>MDM2 protein</th> <th>1.12e+02</th>	S28741	MDM2 protein	1.12e+02
73	6	1.6	81	3 <th>S28741</th> <th>MDM2 protein</th> <th>1.12e+02</th>	S28741	MDM2 protein	1.12e+02
74	6	1.6	95	5 <th>I0801</th> <th>hexosaminidase beta c</th> <th>1.12e+02</th>	I0801	hexosaminidase beta c	1.12e+02
75	6	1.6	125	2 <th>A60426</th> <th>high molecular weight</th> <th>1.12e+02</th>	A60426	high molecular weight	1.12e+02
76	6	1.6	127	2 <th>E64707</th> <th>hypothetical protein</th> <th>1.12e+02</th>	E64707	hypothetical protein	1.12e+02
77	6	1.6	127	3 <th>E64707</th> <th>hypothetical protein</th> <th>1.12e+02</th>	E64707	hypothetical protein	1.12e+02
78	6	1.6	134	5 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
79	6	1.6	134	6 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
80	6	1.6	134	7 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
81	6	1.6	134	8 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
82	6	1.6	134	9 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
83	6	1.6	134	10 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
84	6	1.6	134	11 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
85	6	1.6	134	12 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
86	6	1.6	134	13 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
87	6	1.6	134	14 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
88	6	1.6	134	15 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
89	6	1.6	134	16 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
90	6	1.6	134	17 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
91	6	1.6	134	18 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
92	6	1.6	134	19 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
93	6	1.6	134	20 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
94	6	1.6	134	21 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
95	6	1.6	134	22 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
96	6	1.6	134	23 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
97	6	1.6	134	24 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
98	6	1.6	134	25 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
99	6	1.6	134	26 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02
100	6	1.6	134	27 <th>I12073</th> <th>integrin alpha 1</th> <th>1.12e+02</th>	I12073	integrin alpha 1	1.12e+02

2	11	3.0	353	2	S28787	neuropeptide Y/peptid	1.32e-08
3	10	2.7	360	2 <th>A57160</th> <th>chemokine (C-C) recep</th> <th>2.58e-06</th>	A57160	chemokine (C-C) recep	2.58e-06
4	9	2.5	360	3 <th>A57160</th> <th>chemokine (C-C) recep</th> <th>2.58e-06</th>	A57160	chemokine (C-C) recep	2.58e-06
5	9	2.5	356	2 <th>S42086</th> <th>interleukin-8 recepto</th> <th>3.90e-04</th>	S42086	interleukin-8 recepto	3.90e-04
6	9	2.5	356	3 <th>S42086</th> <th>interleukin-8 recepto</th> <th>3.90e-04</th>	S42086	interleukin-8 recepto	3.90e-04
7	9	2.5	359	2 <th>A48921</th> <th>interleukin-8 recepto</th> <th>3.90e-04</th>	A48921	interleukin-8 recepto	3.90e-04
8	9	2.5	360	2 <th>JC4587</th> <th>chemokine (C-C) recep</th> <th>3.90e-04</th>	JC4587	chemokine (C-C) recep	3.90e-04
9	9	2.5	360	3 <th>JC4587</th> <th>chemokine (C-C) recep</th> <th>3.90e-04</th>	JC4587	chemokine (C-C) recep	3.90e-04
10	9	2.5	362	2 <th>S65631</th> <th>chemokine (C-C) recep</th> <th>3.90e-04</th>	S65631	chemokine (C-C) recep	3.90e-04
11	9	2.5	398	2 <th>S56517</th> <th>mu-opioid receptor</th> <th>3.90e-04</th>	S56517	mu-opioid receptor	3.90e-04
12	9	2.5	398	3 <th>S56517</th> <th>mu-opioid receptor</th> <th>3.90e-04</th>	S56517	mu-opioid receptor	3.90e-04
13	9	2.5	398	4 <th>S56517</th> <th>mu-opioid receptor</th> <th>3.90e-04</th>	S56517	mu-opioid receptor	3.90e-04
14	9	2.5	398	5 <th>S56517</th> <th>mu-opioid receptor</th> <th>3.90e-04</th>	S56517	mu-opioid receptor	3.90e-04
15	8	2.2	394	2 <th>S57333</th> <th>G-protein-coupled rec</th> <th>4.30e-02</th>	S57333	G-protein-coupled rec	4.30e-02
16	8	2.2	372	2 <th>S42592</th> <th>delta opiate receptor</th> <th>4.30e-02</th>	S42592	delta opiate receptor	4.30e-02
17	8	2.2	372	3 <th>S42592</th> <th>delta opiate receptor</th> <th>4.30e-02</th>	S42592	delta opiate receptor	4.30e-02
18	8	2.2	372	4 <th>S42592</th> <th>delta opiate receptor</th> <th>4.30e-02</th>	S42592	delta opiate receptor	4.30e-02
19	8	2.2	372	5 <th>S42592</th> <th>delta opiate receptor</th> <th>4.30e-02</th>	S42592	delta opiate receptor	4.30e-02
20	8	2.2	380	2 <th>JC2434</th> <th>kappa opiate receptor</th> <th>4.30e-02</th>	JC2434	kappa opiate receptor	4.30e-02
21	8	2.2	380	3 <th>JC2434</th> <th>kappa opiate receptor</th> <th>4.30e-02</th>	JC2434	kappa opiate receptor	4.30e-02
22	8	2.2	379	2 <th>JC0894</th> <th>P115 protein - Mycopl</th> <th>4.30e-02</th>	JC0894	P115 protein - Mycopl	4.30e-02
23	8	2.2	379	3 <th>JC0894</th> <th>P115 protein - Mycopl</th> <th>4.30e-02</th>	JC0894	P115 protein - Mycopl	4.30e-02
24	8	2.2	379	4 <th>JC0894</th> <th>P115 protein - Mycopl</th> <th>4.30e-02</th>	JC0894	P115 protein - Mycopl	4.30e-02
25	7	1.9	139	2 <th>S43603</th> <th>RNA binding protein b</th> <th>3.19e+00</th>	S43603	RNA binding protein b	3.19e+00
26	7	1.9	163	2 <th>H64189</th> <th>dead protein - Namop</th> <th>3.19e+00</th>	H64189	dead protein - Namop	3.19e+00
27	7	1.9	323	1 <th>Q08023</th> <th>HRF3 protein - human</th> <th>3.19e+00</th>	Q08023	HRF3 protein - human	3.19e+00
28	7	1.9	350	2 <th>A34443</th> <th>interleukin-8 recepto</th> <th>3.19e+00</th>	A34443	interleukin-8 recepto	3.19e+00
29	7	1.9	350	3 <th>A34443</th> <th>interleukin-8 recepto</th> <th>3.19e+00</th>	A34443	interleukin-8 recepto	3.19e+00
30	7	1.9	352	2 <th>G00048</th> <th>leukocyte-derived sev</th> <th>3.19e+00</th>	G00048	leukocyte-derived sev	3.19e+00
31	7	1.9	352	3 <th>A45747</th> <th>leukocyte-derived sev</th> <th>3.19e+00</th>	A45747	leukocyte-derived sev	3.19e+00
32	7	1.9	354	2 <th>A23669</th> <th>interleukin-8 recepto</th> <th>3.19e+00</th>	A23669	interleukin-8 recepto	3.19e+00
33	7	1.9	355	2 <th>JC0231</th> <th>G-protein-coupled rec</th> <th>3.19e+00</th>	JC0231	G-protein-coupled rec	3.19e+00
34	7	1.9	355	3 <th>JC0231</th> <th>G-protein-coupled rec</th> <th>3.19e+00</th>	JC0231	G-protein-coupled rec	3.19e+00
35	7	1.9	359	2 <th>A22891</th> <th>fix3 protein - Rhizob</th> <th>3.19e+00</th>	A22891	fix3 protein - Rhizob	3.19e+00
36	7	1.9	377	2 <th>P69172</th> <th>hypothetical protein</th> <th>3.19e+00</th>	P69172	hypothetical protein	3.19e+00
37	7	1.9	378	2 <th>S57535</th> <th>lymphocyte-specific G</th> <th>3.19e+00</th>	S57535	lymphocyte-specific G	3.19e+00
38	7	1.9	385	2 <th>S57535</th> <th>lymphocyte-specific G</th> <th>3.19e+00</th>	S57535	lymphocyte-specific G	3.19e+00
39	7	1.9	385	3 <th>S57535</th> <th>lymphocyte-specific G</th> <th>3.19e+00</th>	S57535	lymphocyte-specific G	3.19e+00
40	7	1.9	404	2 <th>JC5784</th> <th>adrenomedullin recept</th> <th>3.19e+00</th>	JC5784	adrenomedullin recept	3.19e+00
41	7	1.9	432	2 <th>S50829</th> <th>alpha 2-adrenoceptor</th> <th>3.19e+00</th>	S50829	alpha 2-adrenoceptor	3.19e+00
42	7	1.9	482	2 <th>S66920</th> <th>probable RNA helicase</th> <th>3.19e+00</th>	S66920	probable RNA helicase	3.19e+00
43	7	1.9	509	2 <th>S64435</th> <th>hypothetical protein</th> <th>3.19e+00</th>	S64435	hypothetical protein	3.19e+00
44	7	1.9	509	3 <th>S64435</th> <th>hypothetical protein</th> <th>3.19e+00</th>	S64435	hypothetical protein	3.19e+00
45	7	1.9	511	2 <th>S44275</th> <th>dopamine receptor pro</th> <th>3.19e+00</th>	S44275	dopamine receptor pro	3.19e+00
46	7	1.9	512	2 <th>S23344</th> <th>hypothetical protein</th> <th>3.19e+00</th>	S23344	hypothetical protein	3.19e+00
47	7	1.9	560	1 <th>RMV77</th> <th>hemagglutinin precurs</th> <th>3.19e+00</th>	RMV77	hemagglutinin precurs	3.19e+00
48	7	1.9	560	2 <th>RMV77</th> <th>hemagglutinin precurs</th> <th>3.19e+00</th>	RMV77	hemagglutinin precurs	3.19e+00
49	7	1.9	563	1 <th>A45519</th> <th>hemagglutinin precurs</th> <th>3.19e+00</th>	A45519	hemagglutinin precurs	3.19e+00
50	7	1.9	563	2 <th>S22083</th> <th>hemagglutinin - influ</th> <th>3.19e+00</th>	S22083	hemagglutinin - influ	3.19e+00
51	7	1.9	563	3 <th>S22083</th> <th>hemagglutinin - influ</th> <th>3.19e+00</th>	S22083	hemagglutinin - influ	3.19e+00
52	7	1.9	563	4 <th>S22083</th> <th>hemagglutinin - influ</th> <th>3.19e+00</th>	S22083	hemagglutinin - influ	3.19e+00

RESULT	ENTRY	TIME	ORGANISM	DATE	SEQUENCE	REVISION	TEXT_CHANGE
1	JC5068	1	G protein-coupled receptor	10-Sep-1997	1	10-Sep-1997	text_change
2	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
3	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
4	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
5	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
6	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
7	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
8	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
9	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
10	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
11	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
12	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
13	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
14	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
15	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
16	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
17	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
18	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
19	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
20	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
21	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
22	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
23	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
24	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
25	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
26	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
27	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
28	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
29	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
30	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
31	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
32	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
33	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
34	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
35	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
36	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
37	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
38	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
39	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
40	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
41	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
42	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
43	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
44	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
45	JC5067	1	Chemokine receptor-like genes	10-Sep-1997	1	10-Sep-1997	text_change
46	JC5067	1	Chemokine receptor-like genes	10-Sep-1997</			



Tue Nov 17 08:55:21 1998

US-08-087-977-10-FBI

```

291 VETVLAERCCNNVPLVTAIGQCFRTFTALIDJADCAVETRTSSQSCAIRTSSIRSQ 350
Db 355 TSTSDNADNALSPTM 369
|||||
|||
Qy 351 TSTSDNADNALSPTM 365
|||||
|||

RESULT 2
ENTRY type complete
TITLE neurotensin 1/peptide YY receptor Y3 - bovine
ORGANISM formal_name Bos primigenius taurus common_name cattle
AUTHOR sequence_revision 17-Apr-1993 text_change
DATE 06-Dec-1996

ACCESSIONS
author S28787
REFERENCE S28787
JOURNAL Mol. Pharmacol. (1991) 40:869-875
TITLE Sequence and expression of a neurotensin Y receptor cDNA.
AUTHOR S28787
ACCESSION residues-type 1-353 #label RIN
#cross-references ENGL:R86739
#keywords
#summary
KEYWORDS
SUMMARY
length 353 molecular_weight 3938 checksum 5502
score 3.04; Score 11; DB 2: Length 353;
Beat Local Similarity 100.04; Pred. No. 1.12e-08;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 LPLVLTLPMAV 97
|||||
|||
Qy 85 LPLVLTLPMAV 95
|||||
|||

RESULT 3
ENTRY type complete
TITLE Chemokine (C-C) receptor 4 - human
ORGANISM formal_name Homo sapiens common_name man
AUTHOR sequence_revision 10-Nov-1995 text_change
DATE 15-Aug-1997
ACCESSIONS
author A57160
REFERENCE A57160
JOURNAL Power, C.A.; Meyer, A.; Nemeth, K.; Bacon, K.B.; Hoogwerf,
AUTHOR A.J.; Proudfoot, A.E.I.; Wells, T.H.C.
TITLE Molecular cloning and functional expression of a novel CC
AUTHOR chemokine receptor cDNA from a human basophilic cell line.
ACCESSION A57160

```

```

preliminary; not compared with conceptual translation
##sequence_type mRNA
##residues 1-360 #label POW
##accession-references GB:U85740
##date
GENE
GDB: CHCHB4
##cross-references
##location G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
FEATURES
       75 .. 97          ##domain transmembrane status predicted label TM1
       97 .. 113         ##domain transmembrane status predicted label TM1
      113 .. 133         ##domain transmembrane status predicted label TM1
      133 .. 175         ##domain transmembrane status predicted label TM1
      175 .. 206         ##domain transmembrane status predicted label TM1
      206 .. 226         ##domain transmembrane status predicted label TM1
      226 .. 268         ##domain transmembrane status predicted label TM1
      268 .. 308         ##domain transmembrane status predicted label TM1
      308 .. 376         ##disulfide_bonds status predicted
      376 .. 410-187     ##binding_site phosphate (Ser) (covalent) (by casein
                        ##kinase II) status predicted (covalent) (by protein
                        ##kinase C) status predicted (covalent) (by protein
                        ##kinase C) status predicted (covalent) (by protein
                        ##predicted)
      410 .. 430         ##binding_site carbohydrate (Asn) (covalent) $status
                        ##predicted$
      430 .. 450         ##length 360 #molecular_weight 41402 #checksum 4713
      450 .. 470         ##Query Match
                        ##Best Local Similarity 100.0% Pos. Max. 252-06 Length 360;
                        ##Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0
SUBJECT
Db 71 RSMOTVILN 80
Qy 69 RSMOTVILN 78
ENTRY
PROSITE 4
ENTRY
TITILE G protein-coupled receptor Z1 - equine herpesvirus 2
ORGANISM Equus caballus
DATE 10-Apr-1996 #sequence_revision 19-Apr-1996 #ext_change
#(Sep-1995)
ACCESSIONS
J04081
REFERENCE
855594
authors Telford, E.A.R.; Watson, M.S.; Alld, H.C.; Perry, J.;
J. Doolittle, A.J. (1995) 249:520-528
#Journal
#Title The DNA sequence of equine herpesvirus 2.

```

Tue Nov 17 08:55:21 1998

US-08-887-977-10. rpr

```

sequence      S51594      preliminary; nucleic acid sequence not shown
status
molecule_type DNA
residues      1-383 #label TEB
#cross-references EMBL:U0695172; FID:0695173
REMARKS
G protein-coupled receptor
SUMMARY
length 383 #molecular-weight 43667 #checksum 1663
Query Match      2.74; Score 10; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 2.58e-06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB      86 GILANILVVI 95
QY      51 GILANILVVI 60

sequence      S42096      #type complete
status
molecule_type mRNA
residues      1-100 #label Rattus norvegicus #common_name Norway rat
#cross-references EMBL:X77797; FID:06754397
REMARKS
G protein-coupled receptor
SUMMARY
length 100 #molecular-weight 40188 #checksum 118
Query Match      2.54; Score 9; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 3.90e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB      310 YAFICQKFR 318
QY      307 YAFICQKFR 315

sequence      A53752      #type complete
status
molecule_type mRNA
residues      1-383 #label Rattus norvegicus #common_name Norway rat
#cross-references EMBL:X77797; FID:06754397
REMARKS
G protein-coupled receptor; transmembrane protein
SUMMARY
length 356 #molecular-weight 40188 #checksum 118
Query Match      2.54; Score 9; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 3.90e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB      310 YAFICQKFR 318
QY      307 YAFICQKFR 315

sequence      A53752      #type complete
status
molecule_type mRNA
residues      1-383 #label Rattus norvegicus #common_name Norway rat
#cross-references EMBL:X77797; FID:06754397
REMARKS
G protein-coupled receptor; transmembrane protein
SUMMARY
length 356 #molecular-weight 40188 #checksum 118
Query Match      2.54; Score 9; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 3.90e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB      310 YAFICQKFR 318
QY      307 YAFICQKFR 315

```

```

10-Sep-1997

ACCESSIONS      A53752
REFERENCE
authors        Prado, G.M.; Thomas, K.M.; Suzuki, H.; Larosa, G.J.;
               J. Biol. Chem. (1994) 269:12997-12998.
journal        J. Biol. Chem. (1994) 269:12997-12998.
title          Molecular characterization of a novel rabbit interleukin-8
               receptor subtype.
accession      A53752
molecule_type cDNA
molecule_type mRNA
#residues      1-358 #label PRA
#crosses-references GB:I24445; NID:q437661; PID:q437662
KEYWORDS
SUMMARY        #length 358 #molecular_weight 46561 #czechsum 2000
               2.59; Score 9; DB 2; Length 358;
               Best Local Similarity 100.0%; Pred. N.30e-04;
               Matches 0; Mismatches 0; Indels 0; Gaps 0;

DB 312 YAFIQMKR 320
QY 307 YAFIQMKR 315
|||||
|||||

RESULT 7
TITLE          A48921
AUTHORS        A48921 type complete
               interleukin-8 type complete
               G-protein coupled receptor Gpcr16 mouse
               #format_name Mus musculus #accession_name house mouse
DATE          20-Feb-1995 #sequence_revision 20-Feb-1995 text_change
A48921: A53677; 149340; 155421; 848909
ACCESSIONS
REFERENCE      Corretti, D.P.; Nelson, N.; Koslosky, C.J.; Morrissey, P.J.;
               Kopelman, N.G.; Gilbert, D.J.; Jenkins, N.A.; Doslik, J.K.;
               Genomics (1993) 18:410-413
               The murine homologue of the human interleukin-8 receptor type
               B maps near the Ity-Lsh-Bcy disease resistance locus.
accession      A48921
molecule_type DNA
#residues      1-359 #label CER
#crosses-references GB:I23637; NID:q435093; PID:q435094
KEYWORDS
SUMMARY        A53677; A53678; A53679; A53680; A53681; A53682; A53683; A53684; A53685; A53686; A53687; A53688; A53689; A53690; A53691; A53692; A53693; A53694; A53695; A53696; A53697; A53698; A53699; A53700; A53701; A53702; A53703; A53704; A53705; A53706; A53707; A53708; A53709; A53710; A53711; A53712; A53713; A53714; A53715; A53716; A53717; A53718; A53719; A53720; A53721; A53722; A53723; A53724; A53725; A53726; A53727; A53728; A53729; A53730; A53731; A53732; A53733; A53734; A53735; A53736; A53737; A53738; A53739; A53740; A53741; A53742; A53743; A53744; A53745; A53746; A53747; A53748; A53749; A53750; A53751; A53752; A53753; A53754; A53755; A53756; A53757; A53758; A53759; A53760; A53761; A53762; A53763; A53764; A53765; A53766; A53767; A53768; A53769; A53770; A53771; A53772; A53773; A53774; A53775; A53776; A53777; A53778; A53779; A53780; A53781; A53782; A53783; A53784; A53785; A53786; A53787; A53788; A53789; A53790; A53791; A53792; A53793; A53794; A53795; A53796; A53797; A53798; A53799; A53800; A53801; A53802; A53803; A53804; A53805; A53806; A53807; A53808; A53809; A53810; A53811; A53812; A53813; A53814; A53815; A53816; A53817; A53818; A53819; A53820; A53821; A53822; A53823; A53824; A53825; A53826; A53827; A53828; A53829; A53830; A53831; A53832; A53833; A53834; A53835; A53836; A53837; A53838; A53839; A53840; A53841; A53842; A53843; A53844; A53845; A53846; A53847; A53848; A53849; A53850; A53851; A53852; A53853; A53854; A53855; A53856; A53857; A53858; A53859; A53860; A53861; A53862; A53863; A53864; A53865; A53866; A53867; A53868; A53869; A53870; A53871; A53872; A53873; A53874; A53875; A53876; A53877; A53878; A53879; A53880; A53881; A53882; A53883; A53884; A53885; A53886; A53887; A53888; A53889; A53890; A53891; A53892; A53893; A53894; A53895; A53896; A53897; A53898; A53899; A53900; A53901; A53902; A53903; A53904; A53905; A53906; A53907; A53908; A53909; A53910; A53911; A53912; A53913; A53914; A53915; A53916; A53917; A53918; A53919; A53920; A53921; A53922; A53923; A53924; A53925; A53926; A53927; A53928; A53929; A53930; A53931; A53932; A53933; A53934; A53935; A53936; A53937; A53938; A53939; A53940; A53941; A53942; A53943; A53944; A53945; A53946; A53947; A53948; A53949; A53950; A53951; A53952; A53953; A53954; A53955; A53956; A53957; A53958; A53959; A53960; A53961; A53962; A53963; A53964; A53965; A53966; A53967; A53968; A53969; A53970; A53971; A53972; A53973; A53974; A53975; A53976; A53977; A53978; A53979; A53980; A53981; A53982; A53983; A53984; A53985; A53986; A53987; A53988; A53989; A53990; A53991; A53992; A53993; A53994; A53995; A53996; A53997; A53998; A53999; A54000; A54001; A54002; A54003; A54004; A54005; A54006; A54007; A54008; A54009; A54010; A54011; A54012; A54013; A54014; A54015; A54016; A54017; A54018; A54019; A54020; A54021; A54022; A54023; A54024; A54025; A54026; A54027; A54028; A54029; A54030; A54031; A54032; A54033; A54034; A54035; A54036; A54037; A54038; A54039; A54040; A54041; A54042; A54043; A54044; A54045; A54046; A54047; A54048; A54049; A54050; A54051; A54052; A54053; A54054; A54055; A54056; A54057; A54058; A54059; A54060; A54061; A54062; A54063; A54064; A54065; A54066; A54067; A54068; A54069; A54070; A54071; A54072; A54073; A54074; A54075; A54076; A54077; A54078; A54079; A54080; A54081; A54082; A54083; A54084; A54085; A54086; A54087; A54088; A54089; A54090; A54091; A54092; A54093; A54094; A54095; A54096; A54097; A54098; A54099; A54100; A54101; A54102; A54103; A54104; A54105; A54106; A54107; A54108; A54109; A54110; A54111; A54112; A54113; A54114; A54115; A54116; A54117; A54118; A54119; A54120; A54121; A54122; A54123; A54124; A54125; A54126; A54127; A54128; A54129; A54130; A54131; A54132; A54133; A54134; A54135; A54136; A54137; A54138; A54139; A54140; A54141; A54142; A54143; A54144; A54145; A54146; A54147; A54148; A54149; A54150; A54151; A54152; A54153; A54154; A54155; A54156; A54157; A54158; A54159; A54160; A54161; A54162; A54163; A54164; A54165; A54166; A54167; A54168; A54169; A5
```











#journal J. Neurosci. (1995) 15:2596-2606  
#title The  $\mu$  opiate receptor: modulation of functional  
desensitization by calcium/calmodulin-dependent protein  
kinase and protein kinase C.  
#cross-references MID:95198115  
#accession 156555  
#status nucleic acid sequence not shown; translated from  
GB/EMBL/DDBJ  
#molecule\_type mRNA  
#residues 1-400 #label RES  
#cross-references GB:U519301; MID:9459831; PID:9459832  
#authors Wang, J.B.; Johnson, P.S.; Perrico, A.M.; Hawkins, A.L.;  
Griffin, C.A.; Uhl, G.R.  
#submission submitted to GenBank, August 1994  
#accession A38997 translated from GB/EMBL/DDBJ  
#molecule\_type mRNA  
#residues 1-50, N', 52-233, V', 235-400 #label WAN  
#cross-references GB:U51119; PID:9452073  
#authors Wang, J.B.; Johnson, P.S.; Perrico, A.M.; Hawkins, A.L.;  
Griffin, C.A.; Uhl, G.R.  
#journal FEBS Lett. (1994) 338:217-222  
#title Human  $\mu$  opiate receptor: cDNA and genomic clones,  
pharmacologic characterization and chromosomal assignment.  
#accession S41072  
#status nucleic acid sequence not shown  
#molecule\_type mRNA  
#residues 51-150, N', 52-400 #label WA2  
#cross-references GB:U51119  
#authors Barr, L.A.; Wansan, E.; Yang, D.  
#journal FEBS Lett. (1994) 354:215-218  
#title Expression of two variants of the human  $\mu$  opiate receptor  
mRNA in SK-N-SH cells and human brain.  
#accession S51137 preliminary  
#molecule\_type mRNA  
#residues 387-400 #label BAR  
#cross-references GB:U51119  
#status GDB:OPRD  
#map\_position 624-645  
#keywords G protein-coupled receptor; glycoprotein; transmembrane  
protein  
FEATURE  
73-96 #domain transmembrane #status predicted #label TM01  
107-132 #domain transmembrane #status predicted #label TM02  
144-165 #domain transmembrane #status predicted #label TM03  
186-208 #domain transmembrane #status predicted #label TM04  
235-257 #domain transmembrane #status predicted #label TM05

283-304 #domain transmembrane #status predicted #label TM06  
321-342 #domain transmembrane #status predicted #label TM07  
371-392 #domain transmembrane #status predicted #label TM08  
#cross-references MID:95198115  
#accession 156555  
#status nucleic acid sequence not shown; translated from  
GB/EMBL/DDBJ  
#molecule\_type mRNA  
#residues 1-400 #label WAN  
#cross-references GB:U51119; PID:9452073  
#authors Wang, J.B.; Johnson, P.S.; Perrico, A.M.; Hawkins, A.L.;  
Griffin, C.A.; Uhl, G.R.  
#submission submitted to GenBank, August 1994  
#accession A38997 translated from GB/EMBL/DDBJ  
#molecule\_type mRNA  
#residues 1-50, N', 52-233, V', 235-400 #label WAN  
#cross-references GB:U51119; PID:9452073  
#authors Wang, J.B.; Johnson, P.S.; Perrico, A.M.; Hawkins, A.L.;  
Griffin, C.A.; Uhl, G.R.  
#journal FEBS Lett. (1994) 338:217-222  
#title Human  $\mu$  opiate receptor: cDNA and genomic clones,  
pharmacologic characterization and chromosomal assignment.  
#accession S41072  
#status nucleic acid sequence not shown  
#molecule\_type mRNA  
#residues 51-150, N', 52-400 #label WA2  
#cross-references GB:U51119  
#authors Barr, L.A.; Wansan, E.; Yang, D.  
#journal FEBS Lett. (1994) 354:215-218  
#title Expression of two variants of the human  $\mu$  opiate receptor  
mRNA in SK-N-SH cells and human brain.  
#accession S51137 preliminary  
#molecule\_type mRNA  
#residues 387-400 #label BAR  
#cross-references GB:U51119  
#status GDB:OPRD  
#map\_position 624-645  
#keywords G protein-coupled receptor; glycoprotein; transmembrane  
protein  
FEATURE  
73-96 #domain transmembrane #status predicted #label TM01  
107-132 #domain transmembrane #status predicted #label TM02  
144-165 #domain transmembrane #status predicted #label TM03  
186-208 #domain transmembrane #status predicted #label TM04  
235-257 #domain transmembrane #status predicted #label TM05

Job time : 70 secs.



























CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR  
CC G-PROTEIN COUPLED RECEPTOR. TRANSMEMBRANE; GLYCOPROTEIN;  
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-DEPENDENT KINASE  
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY  
CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.  
CC  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -2- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC ENBL: X91114; E198175.

DR PROSITE: PS00337: G-PROTEIN RECEPTOR: 1.  
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;  
KW CHEMOKINE;  
FT DOMAIN 1 45 EXTRACELLULAR (POTENTIAL).  
FT TRANSNEM 46 72 1 (POTENTIAL).  
FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).  
FT TRANSNEM 82 112 3 (POTENTIAL).  
FT DOMAIN 113 117 EXTRACELLULAR (POTENTIAL).  
FT TRANSNEM 118 139 3 (POTENTIAL).  
FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).  
FT TRANSNEM 161 180 4 (POTENTIAL).  
FT DOMAIN 181 225 5 (POTENTIAL).  
FT TRANSNEM 226 248 4 (POTENTIAL).  
FT DOMAIN 249 270 6 (POTENTIAL).  
FT TRANSNEM 271 291 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 292 312 2 (POTENTIAL).  
FT TRANSNEM 313 353 4 (POTENTIAL).  
FT DISULFID 116 193 BY SIMILARITY.  
FT CARBOHYD 19 19 POTENTIAL.  
FT NON\_TER 353 353  
SQ SEQUENCE 353 AA: 39919 MW: 4843313 CRC32:

Query Match 2.2%; Score 8; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.03e-03;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 311 YAFIQQNF 318  
|||  
QY 307 YAFIQQNF 314

Search completed: Mon Nov 16 10:35:11 1998  
Job time : 32 secs.







OS	CATOSTOMUS COMMERSONI (WHITE SUCKER).
OC	ETAPIA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA.
CC	OESTRICHITES; ACTINOPTERYGII; CYPRINIFORMES.
RS	(SEQUENCE FROM N.A.)
RC	TISSUE=CHS:
RA	DARLINGTON M.G., HARVEY R.J., GRETER F.R., KREZEKAMP H.J., ZWIERS H.
RA	STIMER T., LEDERIS K., RICHTER D.: J. MOL. EVOL. 37: 68-79, 1993.
CC	-- GENOMIC LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY)
CC	EMBL: Y10904; E1169510; --
DR	PROSITE; PS00137; G-PROTEIN RECEPTOR; 1.
DR	PFAM; PF00011; 7m_1.
DR	NCBI; C38343; G-PROTEIN RECEPTOR, TRANSDUCING, GLYCOPROTEIN.
SQ	SEQUENCE 393 AA; 43232 MW; OUBNCECD CKC32;
	Query Match            2.5% Score 9; DB 13; Length 383;
	Best Local Similarity 100.0%; Pred.No. 648e-05;
	Matches       8; Conservative     0; Mismatches    0; Indels    0; Gaps
DB	326 CLUWVITAP 334
QY	301 CLUWVITAP 309
RESULT	5
ID	O42444
RL	PREFILININ;                  PRT; 366 AA.
DT	01-JAN-1998 (TRENGLAREL 05, CREATED)
DT	01-JAN-1998 (TRENGLAREL 05, LAST SEQUENCE UPDATE)
DT	01-AUG-1998 (TRENGLAREL 07, LAST ANNOTATION UPDATE)
DE	ACTINOTHECIUM BREVISPINUM (MEDITERRANEAN TROUT)
QE	ONCOGENESIS MITOSIS INHIBITOR (SALMO GAIEDNELLI)
OC	ETAPIA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA.
CC	OESTRICHITES; ACTINOPTERYGII; SALMONIFORMES.
RN	(1)
RA	FROM N.A.
RA	ZOO J.P., DANIELS G.D., CUNNINGHAM C., SECOMES C.J.:
RL	SUBMITTED (NOV-1987) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR	EMBL; AM003159; E1193478; --
DR	PFAM; PF00001; 7m_1.
SQ	SEQUENCE 368 AA; 41533 MW; 56D3903B CKC32;
	Query Match            2.2% Score 8; DB 13; Length 368;
	Best Local Similarity 100.0%; Pred.No. 1.30e-02;
	Matches       8; Conservative     0; Mismatches    0; Indels    0; Gaps
DB	134 MILLACITS 141
QY	124 MILLACITS 131

```

EMBL: AF017400; G3098346; -.
KW NEURPEPTIDE.
SQ SEQUENCE 377 AA; 42901 MW; ABID43A CRC32;

Query Match      22%; Score 8; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 1.0e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps

Db 55 LGICGNIL 62
QY 50 LKLGCNIL 57

RESULT          8
ID Q83207       PRELIMINARY;    PRT;   377 AA.
AC AC Q83207.
DT 01-NOV-1996 (TRIMBLABEL_01, CREATED)
DI DI TRIMBLABEL_01, LAST ANNOTATION UPDATE).
DR DR 01-JUN-1998 (TRIMBLABEL_06, LAST ANNOTATION UPDATE)
DS DS PROTEIN-COUPLED RECEPTOR.
OS OS NURINE CYTOMEGALOVIRUS.
SS SS LINEAR; IS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.
SC SC 117MIME; SHELLAN FROM H.A.
RP RP STAIN-SMITH;
RA RA DAVIS-POINTER N.J.; LYNN D.M., VALLY H., SHELLAN G.R., RAWLINSON W.
RD RD JAMAICA, B. 73; 1521-1526(1997).
RE RE J VINO. 73; 1521-1526(1997).
RL RL EMBL: LA1868; GI048723; -.
DR DR PFAM: PF00001; 7cd.1.
SQ SEQUENCE 377 AA; 42205 MW; A01779AC CRC32;

Query Match      22%; Score 8; DB 14; Length 377;
Best Local Similarity 100.0%; Pred. No. 1.0e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps

Db 302 LPNVLXAP 309
QY 302 LPNVLXAP 309

RESULT          9
ID O46074       PRELIMINARY;    PRT;   3345 AA.
AC AC O46074.
DT 04-0674; 1998 (TRIMBLABEL_06, CREATED)
DI DI TRIMBLABEL_06, LAST SEQUENCE UPDATE)
DR DR 01-JUN-1998 (TRIMBLABEL_06, LAST ANNOTATION UPDATE)
DS DS COSMID 3098.
OS OS 3098B.4. MELANOGASTER (FRUIT FLY).
SS SS LINEAR; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
SC SC EUARCHOTER.

```



DT 01-AUG-1988 (TREMBLEL. 07, CREATED  
DZ 01-AUG-1988 (TREMBLEL. 07, LAST ANNOTATION UPDATE)  
DE 01-AUG-1988 (TREMBLEL. 07, LAST ANNOTATION UPDATE)  
DI 01-AUG-1988 (TREMBLEL. 07, LAST ANNOTATION UPDATE)  
DR OREPX PROTEIN.  
DS METRILLOCOCCUS CAPSULATUS: SCOTOBACTERIA, AEROBIC RODS AND COCCI;  
OC METHYLOCOCCACEAE  
OO [1]  
RN SEQUENCE FROM N.A.1309,  
RP TIPPETT, A., KATZ, R.E., HARRIS, J.H.,  
RA TIPPETT, A., JANINE L.; PORALLA K.;  
RE BIOCHIM. BIOPHYS. ACTA 1391:223-232(1986).  
SD ENBL Y09F78; E290E1; -,  
SE SEQUENCE 06 AA, 7198 MW, SC1480AF CWC32,  
Score Match 1.99; Score 7; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 1.64e+00;  
Query Match

[illegible]

2000

[illegible][illegible]